

1972

# The Effect of Syntactic Ambiguity on Judgments of the Comprehensibility and Complexity of Sentences and on Short-Term Memory Capacity.

Carol Cunningham Parr

*Louisiana State University and Agricultural & Mechanical College*

Follow this and additional works at: [https://digitalcommons.lsu.edu/gradschool\\_disstheses](https://digitalcommons.lsu.edu/gradschool_disstheses)

---

## Recommended Citation

Parr, Carol Cunningham, "The Effect of Syntactic Ambiguity on Judgments of the Comprehensibility and Complexity of Sentences and on Short-Term Memory Capacity." (1972). *LSU Historical Dissertations and Theses*. 2303.  
[https://digitalcommons.lsu.edu/gradschool\\_disstheses/2303](https://digitalcommons.lsu.edu/gradschool_disstheses/2303)

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact [gradetd@lsu.edu](mailto:gradetd@lsu.edu).

## INFORMATION TO USERS

This dissertation was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.
2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again -- beginning below the first row and continuing on until complete.
4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

### University Microfilms

300 North Zeeb Road  
Ann Arbor, Michigan 48106

A Xerox Education Company

73-2976

PARR, Carol Cunningham, 1941-

THE EFFECT OF SYNTACTIC AMBIGUITY ON JUDGMENTS  
OF THE COMPREHENSIBILITY AND COMPLEXITY OF  
SENTENCES AND ON SHORT-TERM MEMORY CAPACITY.

The Louisiana State University and Agricultural  
and Mechanical College, Ph.D., 1972  
Language and Literature, linguistics

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THE EFFECT OF SYNTACTIC AMBIGUITY ON JUDGMENTS  
OF THE COMPREHENSIBILITY AND COMPLEXITY  
OF SENTENCES  
AND ON SHORT-TERM MEMORY CAPACITY

A Dissertation

Submitted to the Graduate Faculty of the  
Louisiana State University and  
Agricultural and Mechanical College  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

in

The Program in Linguistics

by  
Carol Cunningham Parr  
B.A., University of Colorado, 1962  
M.A., Louisiana State University, 1970  
August, 1972

PLEASE NOTE:

Some pages may have  
indistinct print.

Filmed as received.

University Microfilms, A Xerox Education Company

## DEDICATION

This work is dedicated to my extraordinary husband, Jim, who good-naturedly devoted so much time and energy to our child and our household in order that I might fulfill my professional ambitions. The usual expressions of thanks to a mate for support and encouragement during a long-term project such as obtaining a graduate degree simply fail to reveal the extent of the contribution he's made to my professional and personal growth.

## ACKNOWLEDGMENT

Dr. J. Donald Ragsdale helped make the writing of this dissertation a learning experience by providing me with the freedom to choose a topic and follow my own instincts in selecting the methods for pursuing it. The encouragement of Dr. Claude L. Shaver and Dr. William W. Evans are also appreciated as is the interest shown by Dr. Mary S. Metz and Dr. Margaret P. Stanley.

Others who deserve my recognition are Dr. Prentiss E. Schilling for his assistance with the statistical analyses of the data, and Mr. Edwin Ryland and Mr. Vernon Villars for making the tape recordings. The cooperation of the high school teachers and Speech Department instructors who allowed me to administer the tests in their classes is also acknowledged. The financial support provided by a Dissertation Year Fellowship has been helpful.

During the course of my studies the exchange of ideas and friendly competition with fellow students Ms. Donna K. Darden and Dr. R. Ed Callary, has been especially stimulating.

## TABLE OF CONTENTS

	PAGE
DEDICATION . . . . .	ii
ACKNOWLEDGMENT . . . . .	iii
LIST OF TABLES . . . . .	vi
LIST OF FIGURES . . . . .	vii
ABSTRACT . . . . .	viii
CHAPTER I. INTRODUCTION AND THEORETICAL CONSIDERATIONS . . . . .	1
Preliminary Remarks	
Data of Linguistics	
Grammatical Analyses of Ambiguity	
CHAPTER II. REVIEW OF EXPERIMENTAL FINDINGS. . . . .	26
Psycholinguistic Studies of Ambiguous Sentences	
Scaling of Linguistic Judgments	
Measurement of Short-Term Memory Load	
CHAPTER III. EXPERIMENTAL DESIGN AND PROCEDURES . .	62
Experiment I . . . . .	64
Hypotheses	
Subjects	
Materials	
Design	
Test Presentation	
Scoring the Tests	
Experiment II. . . . .	75
Hypotheses	
Subjects	
Materials	
Design	
Test Presentation	
Scoring the Tests	
Summary . . . . .	80



CHAPTER IV. RESULTS AND DISCUSSION . . . . .	82
Experiment I . . . . .	83
Reactions to Anomalous Sentences	
Ambiguity Judgment	
Comprehensibility Judgment	
Complexity Judgment	
Summary	
Experiment II . . . . .	.109
Sentence Recall	
Nature of Recall Errors	
Digit Recall	
Summary	
CHAPTER V. SUMMARY AND CONCLUSIONS . . . . .	.129
BIBLIOGRAPHY . . . . .	.136
APPENDIX A. Test Sentences . . . . .	.148
APPENDIX B. Instructions to Subjects in Experiment I . . . . .	.155
APPENDIX C. Answer Sheets for Experiment I . . . . .	.161
APPENDIX D. Instructions to Subjects in Experiment II. . . . .	.164
VITA. . . . .	.165

## LIST OF TABLES

TABLE	PAGE
1. Summary of Experimental Design . . . . .	70
2. Analysis of Variance for Ambiguity Judgment . . . . .	85
3. Analysis of Variance for Degree of Confidence in Judgments of Ambiguity . . . . .	86
4. Percent of Subjects Judging Sentences Ambiguous and Mean Confidence Ratings . . . . .	90
5. Analysis of Variance for Comprehensibility Judgment . . .	97
6. Analysis of Variance for Degree of Confidence in Judgments of Comprehensibility . . . . .	98
7. Analysis of Variance for Complexity Judgment . . . . .	104
8. Analysis of Variance for Degree of Confidence in Judgments of Complexity . . . . .	105
9. Analysis of Variance for Errors Made in Sentence Recall . . . . .	111
10. Nature of Sentence Recall Errors . . . . .	116
11. Analysis of Variance for Digit Recall . . . . .	120

## LIST OF FIGURES

FIGURE	PAGE
1. Necker Cubes . . . . .	7
2. Phrase Structure of (17a) and/or (17b) . . . . .	17
3. Phrase Structure of (17a) and/or (17c) . . . . .	18
4. Type Style Used for Visual Presentation of Test . . . . .	71
5. Mode X Type Interaction Effect on Judgments of Ambiguity . . . . .	89
6. Mode X Type Interaction Effect on Judgments of Comprehensibility . . . . .	100
7. Set X Type Interaction Effect on Digit Recall . . . . .	121

## ABSTRACT

The purpose of this study was to determine the effects of four sources of syntactic ambiguity on judgments of the comprehensibility and complexity of sentences. Whether sentences which are syntactically ambiguous impose a greater load on short-term memory capacity than similar but unambiguous sentences was also investigated.

In Experiment I about 300 high school seniors judged either the ambiguity, the comprehensibility, or the complexity of half of 40 ambiguous-unambiguous pairs of sentences and indicated on a 7-point scale the degree of confidence held in each judgment. Half of the subjects in each group heard tape recordings of the sentences and the other half read them as they were individually projected. It was found that the probability that the ambiguity potential of a sentence will be realized is a function of the source of the ambiguity with the following order from most to least readily recognized: surface structure > deletion > deep structure > case relationships. Furthermore, the mode of presentation of the stimuli used in these tests had no significant differential effect upon the intuitive judgments of the subjects. Results also disclosed that the sentences judged relatively low in comprehensibility and relatively high in complexity were those with relatively high ambiguity indices. Thus, the rating data provide support

for the broad hypothesis that ambiguity contributes to the interpretational density of a linguistic stimulus.

Subjects in Experiment II participated in a short-term memory task which consisted of listening to a tape recording of some of the same sentences, each of which was followed by a string of seven pairs of digits and then recalling the sentence and as many numbers as possible. Under these conditions the unambiguous sentences were not recalled more frequently than the ambiguous sentences nor was digit recall more efficient after the unambiguous than after the ambiguous stimuli. Additionally, the time which elapsed between the presentation and recall of the digits was not inversely related to the number of digits recalled.

These findings suggest that ambiguity complicates the perception of sentences insofar as the dependent measure is rated complexity and that multiple interpretations of ambiguous sentences are not computed in a memory overload situation. The implications of the results of this study are discussed in terms of research methodology and psychological models of human sentence processing.

## CHAPTER I. INTRODUCTION AND THEORETICAL CONSIDERATIONS

For descriptive linguists, particularly those Americans who worked within the framework described by Leonard Bloomfield in Language (1933), the primary data of linguistics consisted of actual utterances of native speakers of various languages. The task of the linguist was viewed as the collection of a corpus of utterances representative of a given language and the subsequent analysis of this corpus into a hierarchy of phonological, morphological, and syntactic units as well as a set of rules outlining permissible arrangements of these units. Informant judgments were elicited only to determine whether similar utterances could be classified as the same or different in meaning, a crucial point in the establishment of a system of autonomous phonemes. This descriptive approach to overt linguistic behavior, often referred to as the structuralist tradition, flourished in this country and its legacies are extensive collections of facts about a diverse array of languages. Then in 1957 Syntactic Structures was published; in this monograph Noam Chomsky argues convincingly in favor of many fundamental changes in the goals and methods of linguistics by pointing out the inadequacy of constituent analyses of the structure of language. In this early work and in numerous subsequent publications (1959, 1961, 1964, 1970a), Chomsky has repeatedly drawn attention to the creative, innovative aspects of language behavior. Hence, he asserts (1964) that one of the goals of a descriptively

adequate grammar is the provision of an account of the principles upon which native speakers rely in the production and comprehension of novel utterances. Since these principles are generally abstract and may not be represented directly in the physical manifestation of utterances, Chomsky argues that the earlier structural grammars which relied entirely on physically measurable, directly observable linguistic data may be observationally adequate insofar as they account for a given set of facts, but are hopelessly impoverished when evaluated in terms of producing insights into the nature of language and human mental processes.

Thus, Chomsky claims, among other things, that observationally adequate grammars fail to reveal, let alone explain, the linguistic intuitions of native speakers, e.g. their ability to recognize paraphrastic relationships between active and passive sentences, their ability to detect ambiguities, their awareness of grammatical deviance. In Aspects of the Theory of Syntax (1965) he writes, "the problem for the grammarian is to construct a description and, where possible, an explanation for the intuition of the native speaker (often himself) (p. 20)." Nowhere does he claim that these intuitions are uniformly reliable, easy to elicit, or even available to the language user; yet he does assert that they constitute the primary data of linguistics and that how well a given grammar accounts for these in-



tutions is one criterion of the adequacy of that grammar, and hence of the conception of grammar and the theory that underlies it.

It is probably valid to criticize theoreticians of transformational grammar for having relied too extensively on introspection and too little on empirical evidence concerning the psycholinguistic abilities of naive speaker-hearers. Of course, obtaining reliable indications of the tacit knowledge users have of their native language does present methodological problems since the implementation of this knowledge is imperfect. However, Chomsky maintains, correctly I think, the following:

it should be clearly recognized that a grammar is not a description of the performance of the speaker, but rather of his linguistic competence, and that a description of competence and a description of performance are different things . . . . Obviously one can find out about competence only by studying performance, but this study must be carried out in devious and clever ways, if any serious result is to be obtained (1970b:43-44).

Essential to understanding the rationale of the research undertaken for this dissertation is the idea that a grammar of a language is a theory of the sentences of that language whose function is to describe the structure of sentences in such a way as to characterize the features used by native speakers in understanding them. Thus, grammars like other scientific theories are expected to make

predictions and different grammars of a given language may, and probably will differ in these predictions. Then these predictions can be empirically tested by comparing them with the linguistic intuitions of native speakers. Most of the predictions made in the present study are derived from the theory of transformational grammar but one subset of the test materials was developed on the basis of a grammar of case relationships.

These preliminary remarks are intended to provide a general outline of the frame of reference within which the present work is conceived. Now before proceeding to a review of empirical investigations of linguistic ambiguity, let us consider some of the ways in which ambiguous sentences have been described and analyzed. It is generally acknowledged that a necessary property of an ambiguous sentence is its potential for sustaining two or more meaningful interpretations regardless of context and situation. This is not to deny that contextual factors affect the disambiguation of sentences but rather comprises an attempt to distinguish between the use of language and language as a system for it is primarily the latter which concerns us here.

Since Kooij's survey (1971:1-7) of the views of Aristotle, Quintilian, Vaugelas, and Bally on ambiguity in natural language is adequate and readily accessible, it would appear otiose for this writer to

pursue the topic in greater detail. Of more immediate interest is Hockett's use (1954) of Necker cubes to provide a visual illustration of ambiguity. From Figure 1 where the drawings are reproduced it can be observed that B is perceived at times like A and at times like C. The linguistic analogy requires that B consist of an ambiguous sentence, A a paraphrase of one of its interpretations, and C a paraphrase of another interpretation. Though this illustration is instructive, there is a danger that it can be carried beyond the limits of its validity since in the perception of ambiguous sentences, awareness of one interpretation does not necessarily preclude awareness of another<sup>1</sup> whereas visually ambiguous figures such as those often used by Gestalt psychologists may not be perceived simultaneously in both ways.

Probably the most common type of ambiguous sentence is that which results from polysemia, i. e. multiple meanings of words. Many puns and quite a few of the riddles which amuse young children illustrate this type of construction as do the following nonhumorous examples:

- (1) The happy children skipped around the block
- (2) Hand him that black pipe

---

<sup>1</sup>Supporting this contention is the fusion effect noted by D. G. MacKay (1966) whereby "both meanings of ambiguous sentences [contribute] to a single integrated interpretation." For instance, the word tank was taken by at least one of his experimental subjects to mean the gasoline receptacle of an armored vehicle.

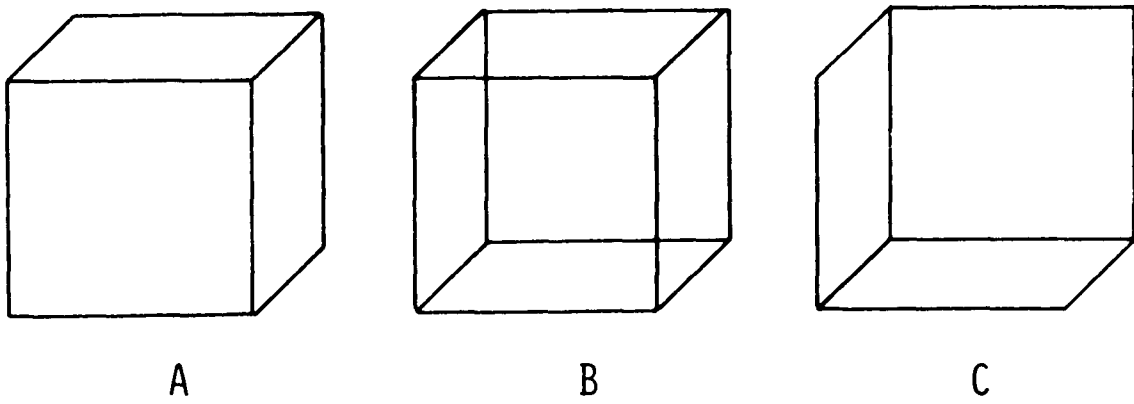


Figure 1. Necker cubes from Hockett (1954:219).

and the multiply ambiguous

(3) The spy took a picture.

In most cases of lexical ambiguity the ambiguous word functions as the same type of constituent in each interpretation although this is not a necessary condition. For instance, in (1) block is a noun which stands in approximately the same grammatical relationship to the verb skipped regardless of whether it is interpreted as referring to a toy, a heavy stand for chopping, an obstruction, or a square formed by the intersection of city streets. Moreover, the occurrence of a polysemous word is, in itself, not sufficient for the sentence in which it is present to be ambiguous since contextual information as well as nonlinguistic information which the listener has about objects and events in the world in some undefined manner constrain his selection of the sense of the lexical item so that inappropriate readings are not often assigned. For instance, if (2) is expanded so that it reads

(2a) The plumber asked me to hand him that black pipe so  
he could repair it

or

(2b) The tobacconist asked me to hand him that black pipe  
so that he could show it to a customer

the associative bonds between the word plumber and the hollow tube interpretation of pipe and between tobacconist and the smoking device interpretation of pipe have the effect of constraining perception of the

meaning of pipe. It is when the linguistic and nonlinguistic context are sufficiently vague that these constraints do not function that sentences are lexically ambiguous. In other words, the more redundant a sentence is, the less likely it is to be ambiguous.<sup>2</sup>

It is also true that the mode of occurrence of certain potentially ambiguous sentences determines the perception of the ambiguity. Consider what your own reactions would be on hearing these sentences:

- (4) That book was /rɛd/
- (5) What a /boʊr/ that is
- (6) Our /sʌn/ is many miles away.

In the case of these and other homonyms the acoustic forms and the respective phonemic transcriptions do not provide the disambiguating information that is available in the orthographic representations. Yet the preceding sentences are unambiguous in the visual mode:

- (4a) That book was red
- (4b) That book was read
- (5a) What a bore that is
- (5b) What a boar that is
- (6a) Our sun is many miles away
- (6b) Our son is many miles away.

On the other hand, some sentences may be lexically ambiguous in the visual mode though unambiguous when heard. Probably everyone has encountered examples similar to this one

- (7) The young girls read very well

---

<sup>2</sup>For a more thorough discussion of the relationships between ambiguity and redundancy refer to Gleason (1965:460-467).

and had to search the surrounding linguistic context for clues as to whether the verb tense was intended to be interpreted as present or past. Moreover, in actual utterances stress may be an indicator of the segmentation of the formatives as has been discussed by Chomsky and Halle (1968), Vanderslice (1970), and Scholes (1971). Certain types of ambiguity can be resolved on the basis of this sort of linguistic information.

In one sense, polysemia can be considered the most elementary type of ambiguity. Quite simply, if at least two distinctly different meanings of a word are reasonably probable within a given linguistic context, then the entire sentence is ambiguous.<sup>3</sup> Whether or not both meanings of the sentence are perceived depends upon that vague "reasonable probability" and, of course, whether or not the hearer knows more than one meaning for the word. Yet in another sense, the detection of non-syntactic ambiguity by native speakers of a language is so complex a phenomenon that in spite of heroic efforts by some philosophers and linguists, inter alia Katz, Fodor, and Weinreich, we are still faced with the situation of having no systematic account of this universal source of ambiguity. So the present study deals exclusively with hypotheses pertaining to various types of

---

<sup>3</sup>Kooij would object to these statements because he makes a distinction between the inherent polysemy of lexical elements and other forms of ambiguity (1971:117-123).

syntactic ambiguity, and the challenging theoretical and experimental problems related to the interpretation of lexical ambiguities are somewhat artificially set aside.

In terms of transformational generative grammar, an ambiguous sentence consists of a single surface structure which can be derived from more than one deep structure configuration.<sup>4</sup> Chomsky (1961) expressed this grammatical principle in the following formal terms:

f (i, j) [which] is the set of structural descriptions of the sentence  $s_i$  that are provided by the grammar  $G_j$  . . . should contain more than one structural description only if the sentence  $s_i$  is ambiguous-- that is, this is a reasonable empirical condition, one of many, on the grammar of a language (p. 120).

Synonymy, the related phenomenon whereby a single deep structure can be represented by more than one surface structure is also quite interesting but not pertinent enough to the analysis of ambiguities to warrant further discussion in this brief survey.<sup>5</sup>

According to Ziff (1965), however, an adequate grammar of a

---

<sup>4</sup>The writer is using the term "deep structure" with full awareness of the current controversy over the nature and even the existence of the level of deep structure as defined by Chomsky (1965). Readers interested in pursuing this topic may refer to Chomsky (1969), McCawley (1968, 1970, 1971), Lakoff (1970), and papers by these and other linguists published in a collection edited by Jacobs and Rosenbaum (1970).

<sup>5</sup>See Gleitman and Gleitman (1970) and Honeck (1971) for reports of some recent psycholinguistic explorations of native speakers' abilities to paraphrase, their judgments of paraphrases, and memory for paraphrases.



language need only provide more than one structural description for all syntactically ambiguous sentence types, i.e. those classes of sentences which are ambiguous by virtue of the arrangement of the words. He concludes his paper by saying "all that one can sensibly ask of a grammar is that it provide a means of discriminating between those sentences that do and those that don't have a syntactic potential for ambiguity (p. 556)." As a consequence, such a grammar cannot distinguish between

(8) I saw the shooting of the apes

in which the syntactic potential is likely to be realized and

(9) I saw the shooting of the elephants

which also has a syntactic potential for ambiguity though it is not likely to be realized.

In Syntactic Structures (1957) Chomsky used the occurrence of different types of syntactic ambiguity to motivate the establishment of phrase structure and transformational structure "as distinct levels of representation for grammatical sequences (p. 85)." His examples of strings requiring at least the phrase structure level for their disambiguation are:

(10) Old men and women<sup>6</sup> . . .

---

<sup>6</sup>Actually this phrase was used earlier by Wells (1947:93ff) who analyzed it in terms of immediate constituents. Dik (1968:228) pointed out that this example and the type of ambiguity it represents have also been discussed by Hockett (1954:217ff and 1961:225-226).

(11) They are flying planes.

These he distinguished from:

(12) I found the boy studying in the library

(13) The shooting of the hunters . . .

(14) John was frightened by the new methods

in which the ambiguities can be resolved only by considering the transformational derivations of the strings and thus determining the fact that the relations between, for instance, shoot and hunters differ depending upon whether the underlying kernel sentence is taken to be

(13a) The hunters shoot

or

(13b) Someone shoots the hunters.

Somewhat similarly, (12) may be interpreted as

(12a) I - found - the boy studying in the library

or

(12b) I - found studying in the library - the boy

just as (14) may be regarded as the passive counterpart of

(14a) New methods frighten John

or as a paraphrase of

(14b) Someone used the new methods to frighten John.

This being the case, it becomes obvious that abstract relations which are not explicit in the physical representation of phrases such as (12), (13), and (14) play a role in the perception of meaning.

Another approach to distinguishing between syntactic ambiguity at the level of surface structure and at the level of deep structure involves the notion of labeled bracketing. This concept is roughly similar to the traditional idea of parsing and draws empirical support from several types of evidence among which are native speakers' intuitions about the strength of grammatical associations between words within a sentence<sup>7</sup> and eye-voice span data such as that acquired in experiments described by Schlesinger (1968). In their discussion of the bracketing notion in general and especially its application to phonology Chomsky and Halle (1968) point out that it is based on the assumption that

the surface structure of a sentence is precisely a proper bracketing of a string of formatives, with the bracketed substrings (the phrases) assigned to categories selected from a certain fixed universal set of categories. The complete string is assigned to the category 'sentence' (S); the other phrases are also assigned to categories that are provided by general linguistic theory, such as the categories 'noun phrase' (NP) and 'verb phrase' (VP) (p. 8).

Now keeping in mind that the disambiguation of sentences in normal situations is a function of pragmatic and semantic constraints, actual resolution of the ambiguity of phrase types with a syntactic potential for ambiguity such as the previously mentioned

---

<sup>7</sup>Levelt (1969, 1970) has developed an interesting technique for quantifying judgments of degrees of relatedness among words and employed it with some success in Dutch and English language experiments.

(10) Old men and women . . .

as well as

(15) Valuable recordings and films . . .

(16) Modern paintings and furnishings . . .

Dik's (1968:236-249)

(17) A cheap bed and breakfast

and an indefinite number of similar phrases can be conceptualized in

terms of labeled bracketing of the strings. For example, contrast

[NP2[A<sup>old</sup> A[NP1men and women]NP1]NP2 with[NP3[NP1old men]

NP1[NP2and women]NP2]NP3.<sup>8</sup> Dik (1968:236ff) argued that such an

analysis cannot, however, account for a third possible interpretation

which arises in cases where there is sufficient internal cohesion be-

tween the coordinated elements to permit modification of the coordina-

tion as a whole. In other words, he contends that the bracketing of a

phrase such as (17) does not enable one to distinguish between

(17a) a bed which is cheap and a breakfast which is cheap

and

(17b) something called a 'bed and breakfast' which is cheap

since the final derived P-marker of (17a) would arbitrarily coincide

with either that of (17b) which is shown in Figure 2 or with that of

---

<sup>8</sup>These illustrations are greatly simplified and the numerical symbols have no theoretical validity but merely serve to assist the reader in identifying pairs of brackets in the schema.

(17c) a bed which is cheap and a breakfast (which need not be cheap)

which is portrayed in Figure 3.

By way of an alternative to bracketing Dik proposed an analysis based upon grammatical functions and explained that in such a system "the different types of coordinative ambiguity can be uniformly accounted for as resulting from differences in the scope of functional relations and from differences in the levels on which the coordinations are situated in the complete structural description (p. 246)."<sup>9</sup>

According to transformational grammar there are as many deep structures and respective derivational histories of a sentence as there are distinctly different syntactic interpretations. Specifically, during the derivation of (17) from the structure underlying (17a) a transformational rule which permits the deletion of the second of two identical formatives (under certain conditions which need not be specified here) has been applied. However, such a rule does not comprise any part of the history of the derivation of (17) from the structures underlying (17b) or (17c). What we seem to have then is a class of sentences which are syntactically ambiguous by virtue of the differential occurrence of deletion transformations during the mapping of deep onto

---

<sup>9</sup>See Dik (1968) for elaboration of the arguments supporting a functional approach and for an outline sketch of such a grammar.

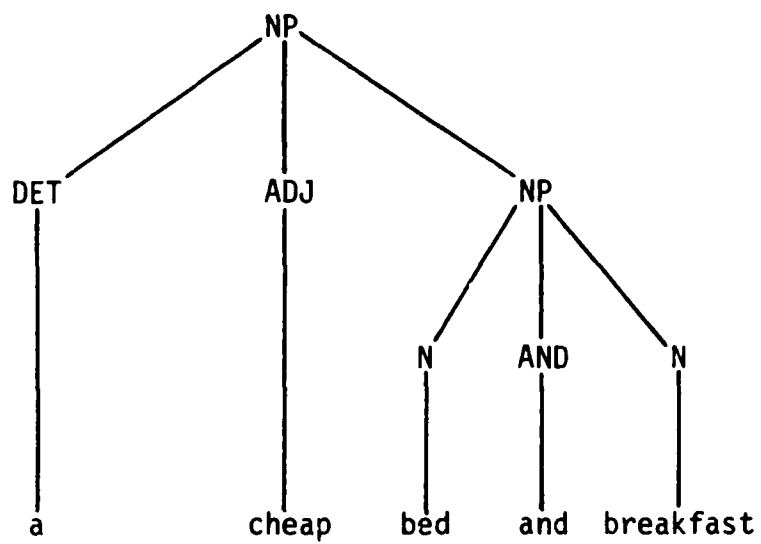


Figure 2. Phrase marker of (17a) and/or (17b)

$[_{NP2} [_{D^a} a] [_{A^a} cheap] [_{NP1} [_{N^1} bed] [_{AND^1} and] [_{N^1} breakfast] ] ]_{NP2}$

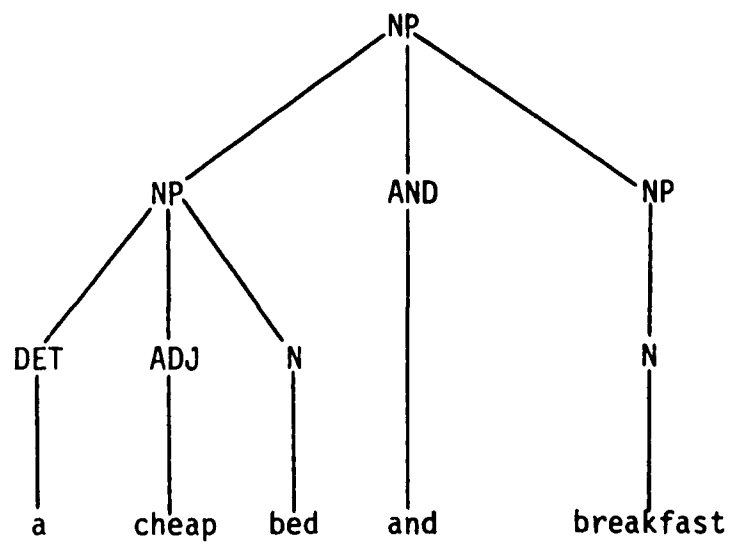


Figure 3. Phrase marker of (17a) and/or (17c)

[NP<sub>3</sub>[NP<sub>1</sub>[D<sup>a</sup>]D[A<sup>cheap</sup>]A[N<sub>1</sub><sup>bed</sup>]N<sub>1</sub>]NP<sub>1</sub>[and]

[NP<sub>2</sub>[N<sub>2</sub><sup>breakfast</sup>]N<sub>2</sub>]NP<sub>2</sub>]NP<sub>3</sub>

surface structures.<sup>10</sup>

Now note what happens when we bracket and label the constituents of the syntactically ambiguous sentence

(18) Growing worms can be amusing:

[S [NP growing worms] NP [VP [AUX can] AUX [V be] V [A amusing] A] VP] S.

Regardless of whether the subject of the sentence is interpreted to mean

(18a) Worms which grow . . .

or

(18b) The growing of worms (by someone) . . .

the same bracketing of the surface structure results. Therefore, since the meaning of the sentence depends not only on the meanings of the constituents, their linear sequence, and their clustering into a hierarchy of units, but also on abstract grammatical relations among them, the ambiguity of this type of sentence must be resolved at a more abstract level. True, were additional constituents indicated above, the labels on the brackets might distinguish between the interpretations but this explanation is not satisfactory since we still cannot describe how the labels are applied without appealing to the function of

---

<sup>10</sup>No directionality is to be inferred from this statement since it is assumed that the same mapping principles operate as part of the linguistic competence of an individual regardless of whether he is functioning as a speaker or a hearer.



the constituents at the level of deep structure. Likewise, the ambiguity of

(19) The special turkey is ready to eat

and

(20) The testing of the psychologist was revealing

cannot be resolved by bracketing of the surface structure but requires a deeper level of analysis.

Thus, according to the theory of transformational generative grammar a distinction can be drawn between several types of syntactically ambiguous sentences. In the work reported here we shall distinguish between the class of sentences ambiguous in surface structure relationships and the class of sentences ambiguous in deep structure relationships by assuming that the application of a labeled bracketing schema can resolve the syntactic ambiguity of the former but not of the latter. Since the terms 'ambiguous in surface structure' and 'ambiguous in deep structure' are used rather consistently throughout the experimental literature to designate these classes, they will also be employed in the remainder of this dissertation but it should be kept in mind by the reader that it is the level at which disambiguation can occur that in fact constitutes the criterion for making the distinction. In other words, even ambiguity that can be resolved through bracketing of the surface structure stems from deep structure differences between

alternative interpretations. In addition, a third class consisting of sentences which are ambiguous by virtue of the application of a deletion rule during the transformational derivation is also established. This group of sentences is referred to by the general term 'deletion' and it tentatively includes sentences in which the ambiguity arises from anaphoric processes such as gapping, sluicing, and clipping.<sup>11</sup>

A different theoretical approach, one which is based upon a linguistic theory of case relationships, will be discussed below for it provides the foundation for establishing the fourth class of sentences used in the tests.<sup>12</sup> First, though, it should be made clear that these four classes by no means exhaust the sources of syntactic ambiguity in English, let alone in other languages. So the taxonomy is rough, the classes neither univocal nor mutually exclusive. Nevertheless it represents a means of proceeding with psycholinguistic investigations and ultimately a step in what one hopes is the right direction.

Grammarians' concerns with case relationships have a long

---

<sup>11</sup>Dingwall (1971) provides definitions and explications of this terminology.

<sup>12</sup>Chomsky (1969a) claims that case systems are not empirically distinguishable from the standard theory of transformational generative grammar since rules mapping case structures onto the deep structures of sentences "can be interpreted as rules of semantic interpretation for these deep structures. Thus one rule (probably universal) will stipulate that for verbs of action, the animate subject may be interpreted as the agent, etc. (p. 9)." Nonetheless, in the present work a separate category is established for this type of sentence.

history which is familiar to anyone even moderately well informed about the development of linguistic inquiry. Charles Fillmore rather recently wrote a stimulating paper entitled "The Case for Case" (1968) which has had the effect of revitalizing the interest of linguists in this approach to analysis. The essence of the position he puts forward is that "the sentence in its basic structure consists of a verb and one or more noun phrases, each associated with the verb in a particular case relationship (1968:21)." He goes on to enumerate some of the cases which frequently occur in natural languages. These are: agentive, objective, dative, instrumental, factitive, locative.

Langendoen is the author of a recent textbook Essentials of English Grammar (1970) which relies heavily on Fillmore's ideas and in this work there are numerous examples of ambiguous case relationships.<sup>13</sup> For instance, the author contends that the sentence

(21) The child was crippled

is ambiguous because crippled can be interpreted as an adjective or as the past participle of the verb cripple, in which case the sentence can be interpreted as the passive counterpart of

(21a) Someone/something crippled the child (p. 56).

Additional examples of sentences which Langendoen finds ambiguous are:

---

<sup>13</sup>Householder (1972) provides us with a brief review.

(22) The hammer struck the nail

(23) The carpenter struck the nail (p. 66).

The possible interpretations can be outlined in terms of case relationships or role structures as follows:

(22a) hammer plays instrument role  
nail plays patient/location role  
 deletion of agent role

(22b) hammer plays patient role  
nail plays patient role.

A paraphrase of interpretation (22a) might be

(22a') Someone/something used the hammer to strike the nail

whereas a paraphrase of interpretation (22b) might read

(22b') The hammer and the nail struck.

Similarly, (23) might mean

(23a) carpenter plays the agent role  
nail plays the patient/location role  
 deletion of instrument role

or, according to Langendoen (1970:66ff),

(23b) carpenter plays the patient role  
nail plays the patient role

the latter interpretation possibly being paraphrased as

(23b') The carpenter and the nail struck.

In these examples it appears that we may be labeling semantic as well as syntactic relationships and thus extending the domain of grammatical analysis into an area which has not yet been shown to be amenable

to transformational description.

Consider the following sentences from the same source (Langendoen, 1970) which, according to this theory, are also ambiguous in underlying case relationships:

(24) John hit the windshield

(25) John slid across the room.

Here it is not clear whether John is in the agent or the patient role.

In a somewhat similar manner,

(26) Ruby is a beautiful soprano

could mean either

(26a) Ruby is a beautiful person who is a soprano

in which instance Ruby would play an essive role, or

(26b) Ruby sings soprano beautifully

whereby Ruby is playing the role of agent. Moreover, the meaning of

(27) John's murderer must have been insane

depends, according to Langendoen, upon whether the identity of John's murderer is or is not known, i. e.

(27a) Whoever was John's murderer must have been insane

or

(27b) The one whom I recognize to be John's murderer must have been insane.

The linguistic intuitions called forth in detecting some of these ambiguities in case relationships are so subtle that one cannot help but

question the validity of some of the distinctions being made. So perhaps it is appropriate to remind ourselves that we are concerned with two aspects of linguistic ambiguity--the potential that exists and the actual realization of this potential. Or one might want to state this problem in terms of the competence of an ideal speaker-hearer and the performance of actual language users.

The remainder of the dissertation is organized as follows. Chapter II consists of a review of experimental literature relating to several areas of interest: psychological parameters in the perception of ambiguous sentences, the elicitation and quantification of native speakers' intuitive judgments about sentences, and a technique for measuring the short-term memory load imposed by a sentence. In Chapter III the hypotheses, design, and procedural details of the present research are described. Chapter IV presents the results of the two experiments and various interpretations of the findings are expressed. Chapter V offers a summary of the work and the main conclusions which can be drawn from it.

## CHAPTER II. REVIEW OF EXPERIMENTAL FINDINGS

The most interesting psychological fact about ambiguous sentences is that not all linguistically permissible interpretations are equally available. Presumably, the commonalities of human experience and information about the nature of the physical world underlie constraints which operate in the comprehension and interpretation of any sentence. Hence, in normal circumstances the probability or likelihood of occurrence of the situations suggested by the various readings determines to a great extent the proportionate availability of the alternative interpretations of ambiguous sentences. For example, in a sentence like

(1) The teasing of the children was cruel

there are no syntactic cues as to whether the nominalized subject is to be perceived in an active or in a passive sense, i.e. whether the children are doing the teasing or being teased. In the absence of a disambiguating context, pragmatic constraints somehow affect the reading a hearer assigns to the sentence. Postal (1964) accurately pointed out that "every piece of possible human knowledge about the world is relevant in the disambiguation of some sentence and thus to its understanding in context . . . [which] shows that theoretically there can be no general theory of the way contexts serve to permit choice of one of several possible interpretations for some sentence . . . (p. 263)."



But rather than abandoning the study of ambiguous sentences, a course of action which might seem to follow from this pessimistic view, transformational generative linguists maintain that native speakers' intuitions about properties of sentences such as well-formedness and susceptibility to more than one structural description comprise the primary data to be accounted for by a grammar of a language. Bever (1970b) asserts that "the fact that linguistic intuitions are subject to the same kind of influences as other types of human judgment does not invalidate many results from linguistic investigations. Many intuitions about sentences appear to be strong enough to resist contextual effects (p. 174)." That fruitful research strategies can be devised within this frame of reference has been demonstrated in the studies of grammaticality, some of which are briefly described in the following review of the literature.

This second chapter can be roughly divided into three sections. Papers reporting experiments in which ambiguous sentences were employed as an independent variable are described first. Then in the second section attention is focused upon several papers whose findings and/or methods seem pertinent to the present investigations of the nature of certain linguistic intuitions. The characteristic shared by the experiments described in the third and final section of this chapter is methodological--in all of them some variation of an ingenious

experimental technique devised by Savin and Perchonock (1965) to investigate the amount of short-term memory "space" occupied by various categories of linguistic stimuli is employed.

### Psycholinguistic Studies of Ambiguous Sentences

A major point of disagreement among investigators has been the description of how a language user proceeds when he encounters an ambiguous sentence. Does he assign it one somewhat tentative interpretation until he receives further information which confirms or refutes it? This characterization is generally referred to as the unitary perception hypothesis though D. G. MacKay (1970) calls it the "garden path" or non-interaction view since the interpretations are assumed to be perceived independently. If it is accurate, can we safely assume that the same strategy is employed in the comprehension of nonambiguous sentences? Or, on the other hand, in normal circumstances does the hearer or reader subconsciously compute all the possible meanings of an ambiguous sentence and become aware of only that interpretation which is most likely as determined by the relevant linguistic and nonlinguistic context (exhaustive computation hypothesis)? A third possibility is that the assignment of any interpretation of an ambiguous sentence is withheld until contextual clues favor one reading over another. MacKay(1966) labels this the oblivion hypothesis. The similarity between the latter two hypotheses is that according to both of

them, two or more interpretations of an ambiguous sentence are processed simultaneously and interact in such a way that one meaning must be suppressed in order for the other meaning to be perceived.

Psycholinguists who have investigated the perception (and the non-perception) of ambiguous sentences have generally employed some measure of the time required to perform certain tasks as the dependent variable. It is obvious that underlying these studies is the assumption that subjects will take more time to perform complex tasks than simple ones; hence the result is an indirect measure of the relative perceptual complexity of certain linguistic stimuli.

MacKay and Bever (1967) found that when subjects were instructed to search for both meanings of ambiguous sentences presented to them visually, the time required to perceive and state the meanings was a function of the type of ambiguity with response times being ordered as follows: deep structure > surface structure > lexical. Moreover, this relative order was maintained when subjects performed the same perception and paraphrase tasks on multiply ambiguous sentences such as

(2) We were surprised at the colonel's appointment.

It was also established in these studies that the bias for either interpretation of an ambiguous sentence, i.e. the percentage of subjects reporting one of the interpretations occurring to them before the other,

entered into a relationship with the perception time such that when there was a marked imbalance in the bias of lexically ambiguous sentences and sentences ambiguous in surface structure the perception times were high. Nevertheless there existed a rather broad range (30%-70% to 70%-30% of the subjects perceiving one interpretation first) through which the median perception time for all types of ambiguities was  $<6$  seconds and  $>4.5$  seconds.

MacKay and Bever also discovered an interesting relation between the bias of a sentence ambiguous at the level of deep structure and the median time required for perception of both interpretations of this type of sentence. Because a similar relation did not hold for surface structure and lexical ambiguities, it was suggested that the interpretations of the latter types of ambiguous sentences are dependent on their relative probabilities and thus coupled, whereas the interpretations of deep structure ambiguities are independent.

When each ambiguous sentence used in the experiment was assigned a complexity value based on the node to terminal node ratio ( $N/TN$ ) of the phrase marker of the terminal string<sup>1</sup> it was found that the higher the  $N/TN$  ratio, i.e. the more structurally complex the sentence, the longer were the perception times for all types of ambiguous sentences investigated. Moreover, "those derived [surface]

---

<sup>1</sup>For a description of the measure see Miller and Chomsky (1963: 480-481).

structure ambiguities involving equal N/TN ratios for both interpretations were perceived faster than those derived [surface] structure ambiguities with unequal N/TN ratios ( $\alpha$  .01), even though there was no bias in the number of subjects seeing the putatively less complex meaning first (p. 197)." However, it should be kept in mind that several investigators (e.g. Fodor and Garrett, 1967; Frank and Osser, 1969; Wang, 1970b) have raised serious questions about the validity of the N/TN ratio as a true indicator of syntactic complexity.

From these findings and the observation that the time needed for perception of two interpretations of multiply ambiguous sentences was significantly longer than for singly ambiguous sentences MacKay and Bever inferred support for the hypothesis that interpretation of an ambiguous word or relationship is held in abeyance until contextual information provides sufficient evidence for favoring one interpretation over another. Thus, according to this view, the longer processing period is a function of the time required to discover the first interpretation of an ambiguous sentence.

MacKay (1966) sought to test several of the hypothetical explanations of how ambiguous sentences are understood and to this end devised a sentence completion test. About half of the sentence fragments he used were singly (2-way) or multiply ambiguous and the remaining fragments consisted of unambiguous versions of the same material.

The college student subjects were instructed to read a sentence fragment such as

(3) Even if I did laugh at the church, I

and then as quickly as possible state an entire sentence, i.e. the fragment and a relevant grammatical completion. The rank order of sentence completion times, stimuli consisting of unambiguous fragments < lexically ambiguous fragments < fragments ambiguous in surface structure < fragments ambiguous in deep structure relationships < multiply ambiguous fragments, confirmed the ordering of perception times reported by MacKay and Bever (1967).

Moreover, MacKay found qualitative differences in the completions; there was a higher incidence of stuttering, misreading, irrelevance, and in one case even laughter in the completions of ambiguous fragments than in the unambiguous versions even though most of the subjects claimed to be unaware of the ambiguities when they were subsequently questioned on this point. In these results MacKay sees support for the oblivion hypothesis, the notion that "ambiguity interferes with the appreciation of a single meaning of ambiguous sentences and that this difficulty is progressively greater for lexical, derived [surface], underlying [deep], and multiple ambiguities in that order (p. 434)."

Corroborating results from a similar sentence completion

experiment were recently reported by MacKay (1970). However, in the more recent tests each ambiguous fragment was preceded by a series of four unambiguous fragments which served to establish an anticipatory set<sup>2</sup> for one type of syntactic structure and thus changed the interpretive bias of the sentence. Nevertheless, the sentence completion times were primarily determined by whether the likely or the unlikely meaning was perceived first. In this paper MacKay outlines a speech perception model which is based on his convictions that comprehension involves both active and passive processes, that sentences can to some extent be processed unconsciously and that the interpretations of an ambiguous sentence can be simultaneously activated and hence interact with each other.

The basic position of Foss, Bever and Silver (1968) who interpreted MacKay's early findings quite differently, is that comprehension is an active, somewhat computational process whereby in the case of ambiguous sentences, a single interpretation is assigned to a string and maintained unless subsequent information contradicts it. They came to this conclusion after studying the results of experiments in which the subjects listened to a sentence and then judged whether a drawing presented to them depicted the meaning of the sentence they

---

<sup>2</sup>J. C. Marshall deserves credit for development of the syntax-setting technique as a psycholinguistic stratagem. See Marshall (1965) for further details.

had heard. The verification times for ambiguous and unambiguous sentences did not differ significantly when the "expected" interpretation of the ambiguous sentence (as determined by independent testing of another group of subjects from the same population) was shown; but if the drawing represented the unlikely interpretation of an ambiguous sentence, the time to verify the picture increased significantly. Foss, Bever and Silver assumed that during this additional time the subject was searching for a second meaning and consequently had not been holding this alternative interpretation in abeyance. Nor did it seem likely to them that the subject had computed a second meaning for the ambiguous sentence and then suppressed it.

Garrett (1970:57) cites an unpublished experiment by Bever in which response times for paraphrasing ambiguous and unambiguous sentences did not differ. Though details of the work are not yet available, it seems that these results provide additional support for Foss, Bever and Silver's view of speech comprehension according to which hearers "typically assign only one immediate interpretation to an ambiguous sentence. Only if that interpretation is found to be incorrect does S reinterpret the sentence . . . . Thus, ambiguity per se does not seem to interfere with understanding the meaning of a sentence (1968:306)."

Foss, et al. do not indicate the bias values for the ambiguous



sentences they used but it is probably accurate to assume that they selected sentences in which one interpretation was much more probable than the other. In contrast, MacKay and Bever (1967) and MacKay (1966) employed ambiguous sentences which they found to fall within the 30-70 to 70-30 range of bias. Consequently they eliminated the sentences with extreme bias from their data. This discrepancy may account for a portion of the seemingly contradictory results. Furthermore, it should also be noted that subjects in the MacKay and Bever (1967) and MacKay (1966) studies read the test sentences or fragments whereas the subjects in experiments conducted by Foss, Bever, and Silver (1968) listened to the stimulus sentences. Whether any of the contradictions in the experimental results can be attributed to this difference in the mode of perception is a question which will be investigated in the present study.

On the other hand, Foss (1970) speculated that the apparent equivocation might result from the fact that MacKay (1966) looked for the effects of ambiguity during the processing of the sentences whereas Foss, Bever and Silver (1968) measured these effects after the sentences were processed. Foss set about testing his hunch by having experimental subjects listen to various combinations of lexically ambiguous sentences, sentences ambiguous in deep structure relationships, and unambiguous sentences constructed by slightly modifying

these sentences so as to eliminate the ambiguity. Responses were timed as each subject either monitored the sentences for a word beginning with /b/, judged whether or not the sentence was ambiguous, or performed both of these tasks. Foss reasoned that "if ambiguity has an effect on spontaneous sentence processing, then longer phoneme monitoring times should be observed during ambiguous than during unambiguous sentences (p. 701)," and this prediction was verified for subjects who participated in just the phoneme monitoring task though not for subjects who monitored for the /b/ phoneme and classified each sentence as to whether or not they found it ambiguous.

Although the type of ambiguity (lexical or deep structure) had no significant effect on the phoneme monitoring response times, it did apparently have an effect on decision times, in that lexical ambiguities were discovered somewhat more rapidly than deep structure ambiguities thus confirming the results of MacKay and Bever (1967).

One of Foss' findings that precludes straight-forward interpretation is that the phoneme monitoring times for subjects who also had to search for ambiguities were slower only for those sentences which they themselves judged to be ambiguous. That is, considerable disagreement existed as to whether or not certain sentences had more than one interpretation. Foss rejected a hypothesis attributing this phenomenon to individual differences in linguistic competence because

the phoneme monitoring speeds of subjects assigned the single task were indeed predicted by his own categorization of ambiguous-unambiguous sentences. Instead, Foss sketches a sentence perception model whereby "the processor must check for the effects of prior information at the point of [potential] ambiguity and this activity itself utilizes some of S's limited analyzing mechanisms . . . . If the context is neutral, then the interpretation of the ambiguity may be determined by some rule of frequency of interpretation or simplicity of resulting structure (p. 705)." So Foss takes the position that ambiguous as well as potentially ambiguous sentences, regardless of whether preceding or succeeding information provides a disambiguating context, are more complex and hence are processed differently than an unambiguous input.

Carey, Mehler, and Bever (1970) employed an experimental technique whereby subjects were "set" to expect a sentence with a particular structural description and then unexpectedly they heard an ambiguous sentence which could be interpreted in a manner compatible or incompatible with the expectation. For instance, one group of subjects listened to a series of five progressive type sentences such as

(4) They are unearthing diamonds

while another group heard a series of five adjectival type sentences

including

(5) They are incoming signals.

The sixth sentence for both groups was

(6) They are visiting sailors.

After hearing each sentence the subject judged it to be true or false with respect to a drawing he had been shown since five seconds before presentation of the sentence. The time which elapsed between the end of the sentence and the subject's utterance of "true" or "false" was measured. Then in a second experiment subjects in whom an expectation for one type of sentence had been established (progressive or adjectival) heard a series of five sentences of the other type and another ambiguous sentence,

(7) They are lecturing doctors.

The same type of true-false judgments with respect to pictorial representations were made for each sentence. Finally, the subjects were asked to paraphrase the ambiguous sentences and in this way indicate which interpretations they had assigned to them.

Although the results from these experiments do not provide unequivocal support for any of the previous hypothetical accounts of the perception of ambiguous sentences, it does appear that the exhaustive computation hypothesis is at odds with the finding that in the first experiment the latency response times of subjects who claimed to have

seen both interpretations of the ambiguous sentence were significantly longer than those of subjects who were not aware of the ambiguities. According to Carey, Mehler and Bever, this and the fact that in experiment I the latencies of subjects who interpreted the ambiguous sentence in accordance with the "set," i.e. assigned it a structure the same as that of the preceding unambiguous sentences, were less than those of subjects who gave a paraphrase indicating an interpretation incompatible with the set suggests that ambiguous sentences can be treated as though they are unambiguous. Further support for the perception hypothesis, i.e. the notion that "those who claimed to have seen only one interpretation in fact did not perceive the other interpretation, and that those who claimed to have seen both interpretations in fact saw both of them (p. 250)" can be inferred from the fact that for those subjects who did not admit to perceiving a second interpretation of the ambiguous sentence, there was no significant difference in the verification response times for the final unambiguous sentence which was fifth in the series and the ambiguous sixth sentence. So, just as Foss, Bever and Silver (1968) showed that strong pragmatic biases can hinder the perception of more than one meaning of an ambiguous sentence as would be predicted by the unitary perception hypothesis, Carey, Mehler, and Bever (1970) have shown that syntactic expectations can have the same effect provided, however, that the subjects' responses

to ambiguity are not distorted by prolonged experience with ambiguous sentences in which case the data become contaminated.

In contrast with the experimental studies of ambiguity described up to this point in which college students were recruited as subjects, Levene (1970) used second, third, fifth, and sixth-grade children. They were tested individually on their understanding and ability to paraphrase sentences which were lexically ambiguous, ambiguous at the level of surface structure, and ambiguous in deep structure relationships. No unambiguous control sentences were employed in the study.

A large piece of equipment designed and constructed for the study made it possible for the experimenter to simultaneously display four drawings; located under each of the four openings for the pictures was a simple switch wired to a timer. The subjects were instructed to read aloud a sentence printed on a card and quickly move the switches beneath the two pictures representing the alternate meanings of the sentence. The time which elapsed between completion of the reading and movement of the switches was measured and was considered indicative of the time used by the subject to comprehend the sentence. Actually, of course, the subjects were searching for ambiguities as did the participants in MacKay and Bever's experiments (1967), and verifying pictorial representations of the meanings they discovered,

a task similar to that requested of subjects by Foss, Bever and Silver (1968). In this type of situation the very presence of the pictures will have the effect that subjects will be made aware of certain interpretations that they might not otherwise notice.

For what Levene calls the "production task" each subject was instructed to read an ambiguous sentence aloud, move one switch as soon as he understood one interpretation, move a second switch when he saw a second interpretation, and then state the two meanings in the order in which they occurred to him. Earlier, MacKay and Bever (1967) employed a similar method with adult subjects; they too measured the latency response times and the number of correct responses. But the data collected in this manner are questionable because according to MacKay (1970) it is wrong to assume "that the time to see a single meaning can be accurately and reliably determined . . . [since] subjects do not know when they have comprehended one meaning of an ambiguous sentence, although they are acutely aware of the point when they have seen both meanings (p. 88)."

Not surprisingly, Levene found that at each grade level the children performed better on the comprehension than the production task and that the older children were more fluent than the younger ones. However, it is intriguing that the pictorial representations of sentences ambiguous at the level of surface structure were correctly

verified significantly more frequently than the drawings representing the meanings of the lexically ambiguous sentences. Moreover, at grade levels 5 and 6 there was no real difference between the verification of sentences ambiguous in deep and surface structure; both types were correctly interpreted more often than lexically ambiguous sentences. Whether these findings which contradict the results of MacKay and Bever (1967) are real or artifacts of the experimental procedure remains to be determined. Now, as Levene suggests it may be that even if the distinction between sentences ambiguous at the levels of deep and surface structure is valid (as was concluded by Mackay, 1966), it simply may not be manifested in the performance of young children.

Yet for the paraphrase or production task (no pictures), correct interpretations were stated more often for the lexically ambiguous sentences than for the other types; moreover, the differences between the number of correct responses to the sentences ambiguous in surface and deep structure relationships were not found to be statistically significant. In fact, for fifth- and sixth-grade subjects the ordering of the mean number of correct responses on the paraphrase task was lexical > deep structure > surface structure.

Also relevant to issues in the perception of ambiguous sentences is the finding of Fodor, Garrett and Bever (1968) to the effect that the greater the variety of deep structure configurations the main verb of a



sentence is capable of entering, the more difficult the sentence is to paraphrase. Though some of the embedded sentences used in their experiment were unintentionally ambiguous, e.g.

- (8) The tactics the general the soldiers admired  
suggested were stupid

the investigators discredited the potential effects of the ambiguity and maintained that "it is the complex lexical structure of the complement verbs that produces difficulty and not some other property of the stimulus sentences (p. 456)." They did acknowledge, however, that syntactic ambiguity might complicate the comprehension of the stimulus sentences.

In a recent review Garrett (1970) argued that ambiguity does in fact complicate the comprehension of sentences in a manner analogous to the lexical complexity of individual formatives. In other words, the presence of structural options increases the computational load on the sentence processing mechanism. This position is quite compatible with experimental findings (e.g. Miller, 1962; Miller, Heise and Lichten, 1951; and Miller and Isard, 1963) to the effect that words in random strings are less intelligible than the same words presented in the context of sentences because syntactic and semantic structures of the sentences constrain the number of interpretive options whereas no such constraints operate in random sequences.

Garrett (1970) also provides an informal description of several

of his own studies of ambiguous sentences. In one experiment in which clicks were superimposed on tape recordings of sentences, it seems that "ambiguity causes an increase in the computational load imposed by a sentence (p. 58)" since polysemous words not disambiguated by prior context had an adverse effect on the subject's ability to accurately locate the click. In a different study in which subjects listened to an ambiguous sentence in one ear while a disambiguating sentence was softly played in the other ear, it was found that although subjects could not report the content of the latter, it effectively determined the interpretation assigned to the ambiguous sentence. Thus, it appears that during the processing of the sentence both interpretations are somehow available to the hearer and this evidence favors a variation of the exhaustive computation hypothesis.

At least brief mention should also be made of Gamlin's finding (1971) that the latency response times of high school subjects to probes<sup>3</sup> tapping ambiguous sentences such as

(9) The starving Indians rode the hungry horses when  
the kind farmer fed them

were greater than latencies to probes tapping "unambiguous" sentences such as

---

<sup>3</sup>The probe-latency technique consists of presenting the experimental subject a sentence and then repeating a word (the probe) from that sentence. The time taken by the subject to respond with the following word from the sentence is measured.

- (10) The starving Indians rode the wild horses when  
the kind farmer fed them.

Inexplicably, subjects with a low short-term memory capacity appeared to be more sensitive to this variable than subjects with greater memory capacity.

By now the magnitude of the difficulties which are encountered when one attempts to sort out the results of these experiments on ambiguity in order to reach definitive conclusions, develop sound hypotheses, and build a theory of speech perception should be apparent. With respect to the test materials, it will be recalled that in some of the studies, e.g. MacKay and Bever (1967) and Levene (1970), no unambiguous controls were employed. Furthermore, linguistic variables which previous studies have revealed as highly significant in the perception and comprehension of sentences have not been consistently controlled; specifically, I am referring to factors such as syntactic complexity, sentence length in terms of number of words or syllables, the frequency of occurrence of the words used in the sentences, the associative relationships that might exist between certain words within the sentences, and the bias favoring one interpretation over another (the latter having been taken into consideration only by MacKay). The mode of presentation of the ambiguous sentences is, of course, another potential source of nonrandom variation.

Now in addition to these factors, consider the diversity of depend-

ent variables which have been measured. Some of them are latency response times in completing sentence fragments, in recognizing ambiguities and paraphrasing their interpretations, in phoneme monitoring, in selecting drawings which depict the meanings of the ambiguous stimulus sentence, in judging whether a picture portrays an interpretation of the sentence, and correct responses in tasks such as these and the subjective location of clicks. And as Garrett (1970:59-60) pointed out, differences in whether the potential effect of an ambiguity was measured during or after sentence processing may be crucial.

The research strategies in most of the studies of ambiguity which I have outlined so far are "psycholinguistic" in the sense of the term elaborated by Bever (1968). According to this view, the description of the formal structure of language provided by transformational generative theory is assumed to be accurate, and by using general behavioral principles to predict certain aspects of linguistic performance the psychological ramifications of the theory are explored. So since it has long been acknowledged in psychological investigations that "a stimulus cannot be perceived in two incompatible ways at the same time (Bever, 1970a:10)," the findings to the effect that ambiguous sentences require more processing time than their unambiguous counterparts hardly seem surprising. Of course it is important to know how linguistic ambiguity interferes with certain perceptual processes. But

from the standpoint of linguistic theory a more direct approach to intuitions about whether or not certain syntactic structures are ambiguous seems desirable.

Thus, the first phase of the research described in subsequent chapters differs in its basic orientation from previous experiments with grammatical ambiguity for here the focus shifts to the formal structure of language as it is systematically reflected in certain intuitions about ambiguous sentences. Moreover, the interaction of several types of linguistic intuition, specifically, judgments of the ambiguity, comprehensibility, and complexity of sentences is scrutinized. Some of the specific questions to which answers are sought in this work are the following: Is it possible to devise a reliable ambiguity index from judgments of the ambiguity of sentences and the degree of confidence held in those judgments? Is such an index of ambiguity sensitive enough to distinguish among several types of syntactic ambiguity? This is a specific instance of the general question of whether there are systematic differences in the realization of the ambiguity potential of sentences representing examples of various sources of ambiguity. Does the mode of presentation of the stimuli (auditory or visual) have a significant effect upon subjects' judgments of the ambiguity, comprehensibility, or complexity of the sentences? In other words, are some ambiguous sentences more

readily perceived as such in one mode than in another? What effects do selected types of syntactic ambiguity have on judgments of the comprehensibility of sentences? What effects do selected types of syntactic ambiguity have on judgments of the complexity of sentences? Does an inverse relationship exist between the judged comprehensibility of a sentence and the judged complexity?

### Scaling of Linguistic Judgments

As we turn now to a brief review of several studies involving subjects' judgments of selected properties of sentences it should be noted that in general the tests of speakers' intuitions toward strings of words varying in degree and kind of semantic and syntactic deviance have been quite straightforward. Consequently the dependent measures often consisted of ratings of the acceptability or grammaticality of the stimuli.

Typical of this direct approach is the work of Maclay and Sleator (1960). In their now classic experiments three groups of college student subjects received instructions to make a binary decision for each string of words they heard with respect to its grammaticality, meaningfulness, or ordinariness. It was found that there was not a perfect fit between the judgments predicted on the basis of linguistic theory and the actual scores obtained for six sets of stimuli (constructed in accordance with selected combinations of the variables  $\pm$  grammatical,

± meaningful, ± ordinary), yet the three types of judgments seem to have been largely independent of each other and the grammaticality judgments, in particular, do distinguish sentences from nonsentences in a manner consistent with generative transformational theory.

Stolz (1969) reported a series of three experiments in which he examined the effect of various types of grammatical deviance on cognitive processing. One group of subjects rated (on a four-point scale) the acceptability of tape recorded sentences, another group performed a short-term memory task patterned after the procedure of Savin and Perchonock (1965), and a third group had to store the sentences in long-term memory before recalling them. In these situations it was observed that violations of selectional restrictions, subcategorization rules, and variations in truth value had more impact on the comprehension of the sentences than on their perception, storage, or recall. Stolz concluded:

if an input sentence is completely processed by the comprehender and results in a well-formed semantic reading, then all subsequent mental manipulations of the reading should depend mainly on such factors as the complexity of the reading (probably a function of the length and syntactic complexity of the original sentence) and the relationship between the content of the reading and the content of LTM [long term memory] (the person's 'internal encyclopedia'). But if the comprehender encounters various structural irregularities in the input, then an incomplete or irrelevant or ill-formed semantic

reading may be produced, causing uniformly more difficult processing in all subsequent operations (p. 217).

In Danks and Glucksberg's experiment (1970) in which a multiple rank ordering technique was used in the elicitation of judgments from native speakers of English as to the grammaticality of some sentences originally used in tests run by Coleman (1965), the high within-subject consistency and between-subject agreement suggested to these investigators that "grammaticalness was scaled as a linear function that was common between subjects (p. 123)." A similar experiment was conducted in which one set of the ungrammatical sentences was interpretable, i.e. in the opinion of the investigators it was rather easy to determine a meaning for them. The higher interpretability of the sentences had a positive effect on the grammaticality ratings, an observation which is congruous with Marks' finding relative to self-embedded clauses (1968) that the more perceptually complex a sentence, the greater the estimates of its ungrammaticalness. Interestingly, Hamilton and Deese (1971) showed that interpretability is also positively correlated with comprehensibility ratings.

In the third phase of their investigation Danks and Glucksberg had four groups of subjects rate sentences on a 10-point scale on the basis of grammaticalness, meaningfulness, familiarity, or ordinari-ness. Principal components analysis of the ratings revealed four



factors which were identified as general comprehensibility, grammaticalness, meaningfulness, and unfamiliarity. Furthermore, it appeared that grammaticalness and meaningfulness could be regarded as independent psychological dimensions.

A number of linguistic factors have been examined in terms of their effect on judgments of sentence comprehensibility. In experiments in which subjects rated sentences (again on a 10-point scale) Danks (1969) found that nonmeaningfulness of the stimuli as defined by diagonalization across five semantically unrelated sentences and ungrammaticalness as defined by anagrammatization, i.e. scrambling of the words within a sentence<sup>4</sup> were reflected in reduced comprehensibility ratings whereas the effects of word frequency, interword association, and syntactic frame, variables which had previously been shown to have an effect on learning (e.g. Rosenberg, 1966; Tulving and Patkau, 1962), were small. As was expected on the basis of the rating data, in direct studies of the comprehension of sentences subjects' latency response times to the nonmeaningful stimuli were greater than latencies to the ungrammatical stimuli. Danks concluded that "either syntactic or semantic deviation [of sentences] lengthened the processing time because of the additional steps involved in providing syntactic

---

<sup>4</sup>The interested reader should refer to Marks and Miller (1964) and Miller and Isard (1963) for a more elaborate description of these techniques and their rationale.

or semantic description. The larger effect of M [meaningfulness] relative to the effect of G [grammaticalness] was expected, since failure to provide an acceptable semantic description may require supplementary syntactic reprocessing before an acceptable semantic description is discovered (p. 695)."

After observing that subjects rated right-branching sentences higher in comprehensibility than center-embedded sentences which were identical in content, Hamilton and Deese (1971) concluded that "the comprehensibility of sentences depends very heavily upon the contiguity, within complex sentences, of the main constituents of each phrase (p. 168)." And although there were some low positive correlations between certain measures of the appropriateness of noun-verb combinations and the comprehensibility rating, the bizarreness or semantic inappropriateness of combinations such as duck dusts, sand feeds, lesson nuzzled had hardly any adverse effect on the ratings, a finding which suggests that grammatical structure is a stronger factor than semantic anomaly in comprehensibility.

Schwartz, Sparkman, and Deese (1971) assert that a "feeling of understanding" underlies judgments of the comprehensibility of sentences and that such a judgment may be made "prior to and independently of interpretation (p. 88)." In other words, they conjecture that a hearer, without actually assigning an interpretation to a stimulus,

may be aware of an existing potential for interpretation. They explored this idea by having ten subjects listen to tape recordings of over 200 sentences and judge the comprehensibility of each. Some of the stimulus variables found to have an effect on the ratings were the number of center-embedded clauses, the number of right branching clauses, semantic anomaly, and logical impossibility of the propositions expressed. Perhaps the most significant contribution of these investigators is their development of a procedure for computing an index of comprehensibility for each sentence. Each experimental subject made a binary judgment indicating whether each sentence was comprehensible or incomprehensible and then rated on a 7-point scale his degree of confidence in that judgment. The score for each stimulus sentence was obtained by multiplying the mean rating scale value by the percentage of comprehensible (C) judgments and subtracting a number equal to the mean rating scale value for incomprehensible (I) judgments multiplied by the percentage of I judgments. Whether in similar fashion the recognition of syntactic ambiguity can be accurately reflected through direct elicitation of binary judgments of ambiguity supplemented with a 7-point self-rating scale on which the subject indicates the degree of confidence he has in his judgment is a question which is explored in the present dissertation.

Participants in Wang's experiments (1970a) were instructed to

rate the comprehensibility of tape-recorded sentences on a 5-point scale and the rank order of the means going from most to least comprehensible sentence types was: normal sentences > anagrams of sentences > semantically anomalous sentences > word lists. This finding supports Danks' observation that nonmeaningful stimuli were rated less comprehensible than ungrammatical stimuli. However, in a recognition task which was also part of the experimental task, the accuracy of judgments according to sentence type was: normal sentences > semantically anomalous sentences > anagrams > word lists, and hence Wang was not able to support her hypothesis that comprehensibility facilitates incidental learning. It is also of interest to note that there were few statistically significant correlations between a subject's rating of the comprehensibility of an item and his subsequent recognition of that item.

Wang (1970b) also investigated the role of syntactic complexity as a determiner of comprehensibility ratings (11-point scale) by computing eight linguistic measures of surface structure complexity for each of seventy-five sentence types. These were the previously mentioned node/terminal node ratio of Miller and Chomsky (1963); the total number of nodes in the phrase marker; the total number of terminal nodes, an indicator of sentence length; a value derived through application of a modification of Yngve's procedure (1960) for calculating

maximum depth; a mean depth value; the number of embedding transformations in the derivational history of the sentence; the number of center-embeddings; and the number of conjoinings. Although the measures based on the mean depth and the number of center embeddings were strongly inversely correlated with the ratings, less than half the variation in these mean comprehensibility ratings could be accounted for by the measures of surface structure complexity. Hence Wang concluded that between sentence differences in deep structures and semantic content also contributed to comprehensibility. The related question of what effect, if any, syntactic ambiguity has on comprehensibility ratings is examined in the first phase of the research undertaken for this dissertation.

### Measurement of Short-Term Memory Load

Before describing the details of those experiments we shall take a quick look at several studies of the effect of some linguistic variables on short-term memory. "Everybody knows that there is a finite span of immediate memory and that for a lot of different kinds of test materials this span is about seven items in length," wrote Miller (1956). Savin and Perchonock (1965) reasoned that if a sentence is encoded separately from its transformational tags, the fewer tags there are to be stored, the more memory space will remain for encoding additional items. Thus, they hypothesized that when a subject attempted to

verbally recall a sentence followed by eight unrelated words, the number of words remembered from the list would vary inversely with the number of transformational tags associated with the sentence. Experimental results confirmed this prediction though it was impossible to avoid the confounding effect of sentence length since transformationally complex sentences, e. g. negatives and passives, are generally longer than their simple, active, affirmative, declarative counterparts. Nevertheless, these investigators did succeed in showing that a psycholinguistic version of the displacement principle could be used to assess certain properties of sentences.

In two close replications of that study Glucksberg and Danks (1969) analyzed errors committed in sentence recall and the number of words recalled from the list following the sentence and, in addition to these variables which had also been measured by their predecessors, they also measured two time periods--that which elapsed between the last item on the word list and the beginning of sentence recall and that which elapsed between the last item on the word list and the beginning of the recall of those words. They summarized their findings by saying that "(a) the error data did not support a transformational model of sentence processing and (b) word recall was, to some extent, a function of syntactic form, but the effects of syntactic form could have been mediated by differential word recall delays, which varied

as a function of syntactic form (p. 215)." The observation that the delay before word recall begins might be significant in determining the pattern of results is pursued in the short-term memory experiment described in the following chapter.

As previously mentioned, Stolz (1969) also used a slight variation of Savin and Perchonock's procedure for investigating the relative degrees of "cognitive impairment" of true, false, contradictory, analytic, scrambled and other sentence types. He found that the word recall data were sensitive to gross differences but could not be used to support the categorization scheme he employed.

Matthews (1968) suspected that word recall in Savin and Perchonock's experimental paradigm was a function of the length of a sentence rather than its transformational complexity and tested his hypothesis by constructing stimulus sentences to vary along both dimensions. He also made several procedural changes. For instance, he asked for written, not verbal recall. And rather than repeatedly using the same five lists of eight words each as Savin and Perchonock had done, Matthews made up different word lists of relatively high frequency words for as many sentences as a subject encountered and in this way eliminated the potential effects of practice in learning the word lists. He also recorded the word lists immediately after each sentence instead of programming a five-second pause at that point and

then elicited the recall responses under two conditions--immediately after the subject heard the last word on the word list and five seconds after the subject heard the end of the list.

Contrary to what one might anticipate, recall of words from the lists was better under the latter condition though little difference between the two conditions was revealed in respect to the number of serious errors committed. The fact that more words were recalled from the list after a five-second delay for each of the 19 sentence categories is puzzling since the relationship is in the opposite direction from that reported by Glucksberg and Danks. In sum, Matthews did not obtain support for the notion that transformational complexity is inversely related to the amount of short-term memory space occupied by a sentence but he did find that sentence length and the presence of adjective qualifiers were factors not to be overlooked.

Epstein (1969) provided further evidence of the susceptibility of the word recall data to variations in output conditions. In order to test his hypothesis that "differences in [word] list recall may reflect differences in reproductive interference and not differences in storage space" he had some of his experimental subjects recall the word list items before they recalled the sentence that had preceded it on the test. Savin and Perchonock's findings were not supported when this particular recall condition was imposed although the systematic differences in



the number of words recalled as a function of sentence type (e. g. interrogative, passive, negative, etc.) that they reported were replicated by Epstein when the subjects were required to recall the sentence before they recalled the words from the list, this being identical with the recall condition of the original experiment.

Still more research based upon Savin and Perchonock's technique was conducted by Wright (1969). There is, however, a discrepancy in her results which arises from the fact that in one experiment the memory load imposed by sentences could be predicted better by Yngve's measure of sentence depth whereas transformational complexity was a better predictor of the memory load in a second experiment. In both studies the recall was written but more control was exercised over semantic variation among the stimuli in the latter than in the former situation so the author concluded that the effect of transformational complexity is greater than the effect of surface structure complexity at least within the limits exercised for the study.<sup>5</sup>

Thus, what was developed by Savin and Perchonock as a technique for indirectly measuring the transformational complexity of sentences, has been shown to have the capability of measuring other properties of

---

<sup>5</sup>Look to Perfetti (1969) for additional experimental data suggesting the inadequacy of the Yngve model and also for a review of some of the theoretical criticisms which have been directed toward some of its fundamental assumptions.

sentences as well. The possibility of determining whether ambiguity is one of those properties motivated the second phase of the present experimental work. Moreover, by establishing whether or not ambiguity affects the memory space occupied by a sentence one could, it seems, test the hypotheses which have been advanced regarding the nature of the psychological processing of ambiguous sentences.

## **CHAPTER III. EXPERIMENTAL DESIGN AND PROCEDURES**

Collecting judgments about sentences from relatively naive informants is a difficult task. Gleitman and Gleitman (1970:8) provide a nice illustration of the problem by speculating what the response of an average taxi driver would be upon being asked, "Is the following a grammatical sentence in your dialect: 'Colorless green ideas sleep furiously?'" In spite of this and other sorts of methodological problems which are faced whenever obtrusive techniques are employed in linguistic research, the results obtained in a direct manner are indispensable since it has repeatedly been demonstrated that one cannot trust a priori assumptions about how native speakers will judge properties of sentences. Such results may be particularly valuable if they can be substantiated through the use of a convergent, less obtrusive procedure.

So with this in mind, the investigator designed two separate experiments, one consisting of the direct elicitation of judgments about ambiguous sentences and their similar but unambiguous counterparts and another which is based upon an indirect measure of properties of the same ambiguous and unambiguous sentences. In the first experiment it is assumed that subjects can (and do) reliably indicate their internal cognitive states in a paper-and-pencil rating task. Moreover, the intuitions about the test sentences are expected to provide a systematic reflection of the perceiver's knowledge of his language, i. e. the

set of rules he has internalized.

The purpose of the first experiment is to determine the effects of four sources of syntactic ambiguity on judgments of the ambiguity, comprehensibility, and complexity of sentences. It will be recalled that in the first chapter there was some discussion of the manner in which some modern grammars analyze the ambiguity potential of certain syntactic constructions. In the present context of concern with linguistic performance, it is the relative probability for the realization of that potential which is under scrutiny.

Here are the research hypotheses which are tested in  
Experiment I:

- (1) The recognition of ambiguous sentences as ambiguous is a function of the syntactic source of the ambiguity with the following order from most to least readily recognized:  
surface structure > deletion > deep structure > case relationships.
- (2) The judgments of a sentence's ambiguity, comprehensibility, and complexity vary depending upon the mode of presentation (aural or visual) of the stimuli.
- (3) Syntactic ambiguity affects native speakers' intuitions about sentences in such a way that within each group of sentences the ambiguous type sentences are judged lower in compre-

hensibility than their unambiguous counterparts.

- (4) Within each group of sentences the complexity rating for ambiguous type sentences is greater than the complexity rating for unambiguous type sentences.
- (5) The sentences judged relatively low in comprehensibility and relatively high in complexity are those with a relatively high probability for realization of the ambiguity potential.

### Subjects

The investigator made arrangements with appropriate educational administrators for testing students in four randomly selected senior English classes at each of three high schools in East Baton Rouge Parish in January of 1972. However, because of a community crisis, attendance was very low at one of the high schools on the day scheduled for testing, so an additional class was tested and in several instances students from more than one section were brought together for a common testing session. With these few exceptions, students sat at the desks they normally occupy in their own classrooms while they took the test.

According to teachers, twelfth grade students who enroll in the optional English course are primarily those who will continue their education at the college level. A total of 312 students took the tests. Males comprised 51% of the sample, females 49%. Most of the

subjects were 17 or 18 years of age with several (N= 9) 16-year olds and one 19-year old randomly distributed throughout the sections. Data from subjects who misunderstood the instructions and failed to indicate both a judgment and the degree of confidence in it (N= 16) or were not native speakers of English (N= 1) were not analyzed.

### Materials

The forty pairs of sentences, ten anomalous strings of words, and five practice items used in Experiment I are shown in Appendix A. Ten sentences were assembled representing each of these four potential sources of syntactic ambiguity: surface structure relationships; deletion; deep structure; case relationships. Most of the sentences are original, although some, particularly those ambiguous in case relationships, were borrowed from examples in the literature.

For each of these forty sentences an unambiguous counterpart was constructed by very slightly changing the choice or the arrangement of the words. It is true, of course, that no sentence is absolutely unambiguous since an inherent characteristic of human language is vagueness or what Waisman has referred to as "open texture." Hence, a clever reader may be able to devise more than one interpretation for the so-called unambiguous members of the pairs of sentences. Nevertheless the distinction appears to have sufficient validity for the purposes of the present experiment.

One set of anomalous sentences ( $N=5$ ) was constructed by selecting words along the diagonals of a matrix of seven consecutive ambiguous or unambiguous sentences within each category of ambiguity source and then scrambling the sequence of each resulting string. The remaining anomalous sentences ( $N=5$ ) were constructed by diagonalizing across the four groups of sentences to obtain the words and then randomly rearranging their order. The purpose of including the anomalous sentences in the tests was to sustain subjects' interest in the experimental task.

In addition to these 50 items, five fillers were used at the beginning of every test to absorb warm-up effects. Like the anomalous sequences, these practice items, which are also shown in Appendix A, were not scored.

In order to control the possible effect of sentence length on the judgments made by the subjects, the total length of each sentence was limited to  $7 \pm 2$  words. This constraint is somewhat artificial, given the fact that function words are more predictable than content words and thus contribute less to the total meaning of a sentence. Yet crude as the restriction may seem, it undoubtedly reduced inter-sentence variation along this dimension to a tolerable level.

Moreover, so as to establish that the size of an individual's vocabulary would not limit his performance on the test, only words



occurring ten or more times in the million word corpus of contemporary American English compiled by Kucera and Francis (1967) were employed in the stimulus items. This restriction on the lexicon made the construction of interesting sentences very difficult and, in retrospect, it seems that it may have been too severe.

### Design

A randomized split-plot experimental design was used. Because of the distortion which might result if subjects judged the similar but ambiguous and unambiguous members of each pair, it was necessary to administer two forms of each test. These forms are labeled A and B in the following commentary and the same letters are used to designate which sentences comprise each test in Appendix A. Consulting these lists shows that Form A consists of five ambiguous and five unambiguous sentences from each of the four groups whereas Form B consists of the twenty unambiguous counterparts of the A sentences and the remaining twenty ambiguous sentences. The first five items on Forms A and B are the practice sentences. A table of random numbers was used to randomize the sequence of the remaining 50 items (20 ambiguous, 20 unambiguous, and 10 anomalous). The order of items was identical for both forms. For example, the thirty-seventh item on Form A is

- (1) These friends are the ones to help

One set of anomalous sentences ( $N=5$ ) was constructed by selecting words along the diagonals of a matrix of seven consecutive ambiguous or unambiguous sentences within each category of ambiguity source and then scrambling the sequence of each resulting string. The remaining anomalous sentences ( $N=5$ ) were constructed by diagonalizing across the four groups of sentences to obtain the words and then randomly rearranging their order. The purpose of including the anomalous sentences in the tests was to sustain subjects' interest in the experimental task.

In addition to these 50 items, five fillers were used at the beginning of every test to absorb warm-up effects. Like the anomalous sequences, these practice items, which are also shown in Appendix A, were not scored.

In order to control the possible effect of sentence length on the judgments made by the subjects, the total length of each sentence was limited to  $7 \pm 2$  words. This constraint is somewhat artificial, given the fact that function words are more predictable than content words and thus contribute less to the total meaning of a sentence. Yet crude as the restriction may seem, it undoubtedly reduced inter-sentence variation along this dimension to a tolerable level.

Moreover, so as to establish that the size of an individual's vocabulary would not limit his performance on the test, only words

occurring ten or more times in the million word corpus of contemporary American English compiled by Kućera and Francis (1967) were employed in the stimulus items. This restriction on the lexicon made the construction of interesting sentences very difficult and, in retrospect, it seems that it may have been too severe.

### Design

A randomized split-plot experimental design was used. Because of the distortion which might be introduced by the judgment of the similar but ambiguous and unambiguous sentences, it was necessary to administer two forms. The first form is labeled A and the second form is labeled B in the following comparison. The first five items on Form A are used to designate which sentences are ambiguous and which are unambiguous. Appendix A. Consulting these lists shows that there are five ambiguous and five unambiguous sentences from each of the four groups whereas Form B consists of the twenty unambiguous counterparts of the A sentences and the remaining twenty ambiguous sentences. The first five items on Forms A and B are the practice sentences. A table of random numbers was used to randomize the sequence of the remaining 50 items (20 ambiguous, 20 unambiguous, and 10 anomalous). The order of items was identical for both forms. For example, the thirty-seventh item on Form A is

- (1) These friends are the ones to help

and the thirty-seventh item on Form B is the unambiguous counterpart,

(2) These friends are the ones who help.

Four class sections were randomly assigned to each of the three task sets, i. e. judging the ambiguity, comprehensibility, or complexity of the stimuli. Ideally, with 25 students in each section, data from 100 subjects would be included in each set. The fact that such an ideal situation is rarely, if ever, encountered in behavioral research is not alarming though it does present a few minor difficulties in the analysis of the experimental results. Two of the four sections in each task set took Form A of the test and two took Form B. Each form was presented in the visual and the aural modes resulting in a design which is summarized in Table 1.

### Test Presentations

For the visual presentation, each of the 55 items and its corresponding serial number was typed in IBM orator style, upper case, without punctuation marks other than the apostrophe. Figure 4 provides some examples of the appearance of the stimuli. Acetate transparencies of the two forms of the test were prepared.

At the beginning of each test session the experimenter introduced herself and distributed test booklets which consisted of a personal data sheet, a page of instructions, and five pages for recording responses. Subjects were reassured that their performance on the

**TABLE 1**  
**Summary of Experimental Design**  
**Showing Unequal Numbers of Subjects in Groups**

<u>Task Set</u>	<u>Form A</u>		<u>Form B</u>		<u>TOTAL</u>
	<u>Aural Mode</u>	<u>Visual Mode</u>	<u>Aural Mode</u>	<u>Visual Mode</u>	
Ambiguity	23 A	29 B & C	32 C	22 A	106
Comprehensibility	40 B	16 C	23 A	28 B	107
Complexity	26 C	22 A	17 B	17 C	82

The letters in the body of the table designate the high schools where the tests were administered: A, Robert E. Lee High School; B, Tara High School; C, Baton Rouge High School.

1. WE FOUND WHAT WE WERE LOOKING FOR

6. JOE WATCHES THE FISH AND HIS CAT

15. THE OLD CHAIR SLID ACROSS THE ROOM

19. IT'S GOOD FOR THEM TO WIN THIS

29. THE HEAVY HAMMER STRUCK THE NAIL

41. POLICE PURSUED THE MAN WITH A CAR

Figure 4. Type Style Used for Visual Presentation of Test

test would not affect their grades, could not, in fact, since their names were not solicited. They were, however, asked to indicate their age, sex, and native language.

Obviously, the instructions, reproduced in Appendix B, differed depending upon whether the class was judging the ambiguity, comprehensibility, or complexity of the sentences and whether the mode of presentation was aural or visual. Judgments were trichotomous--ambiguous, unambiguous, meaningless--in the first instance and dichotomous--comprehensible, incomprehensible or complex, not complex--in the latter sets. It should be noted that in addition to the judgment itself, the degree of confidence held in it was also to be indicated on the answer sheet for each item.

The experimenter read the instructions aloud to the entire class. Seldom did questions regarding the nature of the experimental task arise and these were answered by rereading appropriate passages from the instruction sheet. Questions regarding the purpose of the test were put off until completion of the test. The experimenter began the visual test by positioning a movie projection screen at the front of the classroom and turning off the lights. A 3M overhead projector was used to display the stimuli, the acetate transparencies being shielded in such a way that no more than one sentence was visible at a time. The experimenter called out the item number when each string first appeared

on the screen and after a 20-second projection period moved the shield so as to uncover the next item. This was a long interval between sentences, especially in light of Schwartz, Sparkman and Deese's recommendation (1970) that the inter-sentence interval be held to less than the three seconds which they used in their experiments. However, since MacKay (1966) and MacKay and Bever (1967) showed that subjects normally required considerably more time than this in order to perceive more than one interpretation of an ambiguous sentence, the longer interval seemed necessary. After the students finished responding to all 55 items they were asked to write a few brief comments as to whether they noticed anything unusual about the sentences they had seen.

A total of about 40 minutes was required for administration of the test and since the class periods were 45-50 minutes long, some time remained for debriefing the students and satisfying their curiosity as to the purpose of the experiment and their role in it.

The format for the auditory presentation of the test was the same as that just described for the visual presentation. In this case, however, a Uher 4000 Report-L monophonic tape recorder was used to present the sentences. Earlier, an adult male native speaker of English with no obvious regional dialect had made a tape recording of forms A and B of the test. Reasonable care was exercised with the



ambiguous sentences so that patterns of stress and juncture neither excluded nor favored one interpretation over another. The recordings were made in a professional sound studio on an Ampex tape recorder with an RCA microphone Type BK-5B MI-11010-A on Sony PR-200 polyester tape at a speed of  $7\frac{1}{2}$  ips. The serial order of items was the same in the visual and auditory tests as was the inter-sentence interval of 20 seconds.

#### Scoring the Tests

Copies of the answer sheets are shown in Appendix C. Each item of interest, i. e. the forty pairs of sentences, was assigned a code according to the sentence group (surface structure, deletion, deep structure, case relationships) and type (ambiguous or unambiguous). The judgment and degree of confidence provided by each subject for each item on the test were recorded separately. In the ambiguity task set judgments were scored in terms of whether or not they agreed with the theoretical classification. That is to say, a judgment of ambiguous for an ambiguous sentence was scored as "correct," whereas a judgment of unambiguous or meaningless was scored as "incorrect." Responses were not recorded for those few random instances where subjects neglected to indicate both types of information. No satisfactory method was devised for quantifying the subjects' comments so they were not subjected to statistical analysis.

Some readers may, at this point, wish to proceed to Chapter IV where the results of this first experiment are presented and discussed.

Three variables are measured in Experiment II: the recall of a sentence (correct or incorrect), the number of decimal digits recalled after each sentence, and the perception or non-perception of ambiguity.

The research hypotheses are as follows:

- (1) Unambiguous sentences are recalled more frequently than ambiguous sentences.
- (2) Within each group of sentences, the number of errors in the recall of ambiguous sentences is greater than the number of errors in the recall of unambiguous sentences.
- (3) Within each group of sentences (across all subjects), the mean number of digits recalled after the unambiguous sentences is greater than the mean number of digits recalled after the ambiguous sentences.
- (4) Across all four groups of sentences and across all subjects (i. e. within each pair of sentences), the mean number of digits recalled after the unambiguous sentences is greater than the mean number of digits recalled after the ambiguous sentences.
- (5) Across all four groups of sentences and within each subject, the mean number of digits recalled after unambiguous

sentences is greater than the mean number of digits recalled after sentences perceived as being ambiguous.

- (6) The time which elapses between the presentation and recall of digits as measured by the number of alphabetic characters transcribed per sentence is inversely related to digit recall.

### Subjects

The subjects for this experiment consisted of undergraduate students enrolled in four randomly selected sections of the Speech 2 course, Voice and Articulation, at Louisiana State University in the spring of 1972. A total of 82 subjects, 75 females and 7 males, participated in the experiment.

### Materials

It will be remembered that 10 pairs of sentences, controlled for overall length and word frequency, representing four sources of syntactic ambiguity were used in Experiment I. Because of strict limitations on the time period available for administering the short-term memory tests, only 8 of the 10 pairs of sentences from each of the four groups were incorporated into the second experiment, resulting in a total of 32 items of interest. The anomalous strings were not used but five filler sentences, four at the beginning and one at the end, were used to absorb warm-up and termination effects, respectively.

The order of the remaining items was the same as in the first experiment.

Whereas the recall of strings of 8 English words was measured in previous research involving the effects of certain properties of sentences on the amount of space remaining in short-term memory, pairs of decimal digits were used in the present investigation. This change was made for several reasons, one being the possibility that English words are not equally difficult to remember. For example, this writer finds color names much easier to remember than the names of vehicles (truck, bus, car, etc), these being two of the categories previously employed. Another reason is that associations which undoubtedly would have the effect of enhancing recall are not so readily formed among digits as among words. The subjects were, therefore, told that the order of the first members of the digit pairs was fixed so that the first pair ranged from 20 to 29, the second pair from 30 to 39, the third pair from 40 to 49, and so forth. The second members of the pairs were selected from a table of random numbers.

### Design

Once again the problem of administering the test to groups of unequal numbers of students was encountered. And as in the first experiment, it was necessary to administer an A and a B form of the test, 42 subjects comprising the set taking the former and 40 subjects

taking the latter. Sentence group and type were the independent variables, and the effects of different modes of presentation were not examined.

### Test Presentation

A copy of the instructions read to each class of students can be found in Appendix D. Subjects were provided with 4 X 6 inch pads of plain paper for recording their responses, one test item, i.e. sentence plus series of digits, per page.

The tapes which were played to the groups on Uher 4000 Report-L equipment had been made by an adult male native speaker of English on Ampex recording equipment at a tape speed of  $7\frac{1}{2}$  ips.

The format was as follows: a sentence uttered at a natural rate of speed followed by a very brief pause, then seven pairs of decimal digits in the order just described, and the command "begin." A 40-second interval was programmed for the subjects to record their responses. The experimenter monitored the session to ensure that subjects recalled the entire sentence before the digits, though the digits themselves could be recalled in any order. Two rest breaks of several minutes duration were provided during the test session. Upon completion of the 37 items the subjects were asked to go through their answer pads and underline all the sentences which they judged to be ambiguous. Thus, not until the short-term memory task had

been completed were the subjects told that some of the stimuli were ambiguous.

### Scoring the Tests

Criteria were established in such a way that sentences could be scored as correctly recalled even if one of the following types of errors were made: the substitution of one homonym or near-synonym provided that the ambiguity potential not be affected; the permutation of co-ordinated elements; a change of the verb tense; a misspelled word. Data from subjects who committed more than ten major errors in recalling the 32 sentences of interest were excluded from analysis (N= 11). Of the remaining subjects, 32 had taken Form A of the test, 39 Form B. If a subject recalled a sentence within the established limits of accuracy, then the number of pairs of digits recalled was also scored as was his perception or non-perception of the ambiguity.

In addition to coding information as to the sentence group (surface structure, deletion, deep structure, case relationships) and type (ambiguous or unambiguous), the investigator also recorded the number of alphabetic characters in each sentence. This information provides an indication of sentence length, and, as a consequence, the relative amount of time which passes between the presentation and the recall of the digit pairs. Thus, the letters are not in any way equated with encoding or decoding units. Indeed, such an assumption would be

quite unfounded in light of the considerable amount of evidence which supports the contention that the encoding units or "chunks" consist not of letters (or phonemes), or even single words, but rather of groups of words.

Summarizing briefly, in Experiment I about 300 high school seniors in Baton Rouge judged either the ambiguity, the comprehensibility, or the complexity of half of 40 pairs of sentences and indicated on a 7-point scale the degree of confidence they had in each judgment. Half of the subjects in each group heard tape recordings of the sentences and the other half read them as they were individually projected. Thus, the between subject variables are the nature of the rating task and the mode of presentation of the stimuli. The within subject variables are the sentence types: unambiguous, syntactically ambiguous in surface structure relationships, syntactically ambiguous as a result of the deletion of elements from the surface structure, syntactically ambiguous in deep structure relationships, and ambiguous in case relationships.

University undergraduate subjects in Experiment II participated in a short-term memory task which consisted of listening to a tape recording of some of the same sentences, each of which was followed by a string of seven pairs of digits and then recalling the sentence and as many numbers as possible. Guiding this experiment was the

general hypothesis that ambiguous sentences are more difficult to process, store, and recall than unambiguous sentences. This being the case, the frequency of recall errors made by subjects should be higher for ambiguous than for unambiguous sentences and, furthermore, because of their greater interpretational density, the ambiguous sentences should tax a subject's short-term memory capacity to a greater extent than the unambiguous sentences. Therefore, it was predicted that the mean number of unrelated items (digits) recalled after ambiguous sentences is less than the mean number of items recalled from presentation in a list after unambiguous sentences.



## **CHAPTER IV. RESULTS AND DISCUSSION**

### Experiment I

The responses of the subjects in the first experiment were sorted into three task sets according to the nature of the judgment and were analyzed separately. In addition, separate statistical tests were run within each set for the judgment itself and for the degree of confidence expressed in it. The results of these tests are presented in the pages that follow, but first a few brief comments on the nature of the subjects' reaction to the anomalous sentences seem in order.

Although judgments of the deviant items were elicited, the responses were not recorded or analyzed. The investigator did, however, unobtrusively note that after the presentation of an anomalous item, the noise level in the classroom increased considerably. This was due to the students' coughing, sighing, mumbling, throat clearing, talking, pencil dropping, laughing, and just general fidgeting. In these ways subjects manifested their anxiety over the oddity of the stimuli. These same reactions also occurred, but to a much lesser extent, after ambiguous items such as

(23B) The foolish boy gave her cat food  
and

(34B) Lucy likes money more than her husband.

### Ambiguity Judgments

The subjects' judgments of the ambiguous sentences yielded data which could be regarded as estimates of the probability that the ambiguity potential of the test sentences would be realized. Or the judgments could also be viewed in terms of "correctness," i. e. agreement of the subjects' perceptions with the theoretical classification scheme of the investigator. It will be recalled that subjects in this task set chose one of the three terms--ambiguous, unambiguous, meaningless--to describe the stimulus. When an ambiguous sentence was judged as such, the response was scored as "right," but when it was judged unambiguous or meaningless it was scored "wrong." Likewise, an unambiguous sentence judged to be so was scored "right," but judgments that it was ambiguous and meaningless were lumped together. Data coded in this manner were subjected to analyses of variance and the results of these tests are shown in Tables 2 and 3.

The tables disclose that the main effects of sentence type and group are statistically significant for both the judgment ( $F = 100.58$ ,  $p < .01$ ;  $F = 11.07$ ,  $p < .01$ ) and the degree of confidence ( $F = 16.89$ ,  $p < .01$ ;  $F = 5.04$ ,  $p < .01$ ) in the judgments. The values of these ratios are derived from a comparison of the means for the ambiguous (45.76% correct, 6.42 mean confidence) with the unambiguous sentences (60.36% correct, 6.31 mean confidence) and of the means of the

**TABLE 2**  
**Analysis of Variance for Ambiguity Judgment**  
**(N = 104)**

<u>Source of Variation</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>	<u>F</u>
Mode	1	.32	.32	<1
Subjects (Mode)	104	33.87	.33	
Type	1	22.53	22.53	100.58**
Mode X Type	1	1.88	1.88	8.39**
Group	3	7.43	2.48	11.07**
Sentences (Group)	36	61.02	1.69	7.59**
Mode X Group	3	.46	.15	<1
Type X Group	3	34.52	11.51	51.34**
Mode X Type X Group	3	.07	.02	<1
Error	4076	911.90	.22	

\*\*  $p < .01$

**TABLE 3**  
**Analysis of Variance for Degree of Confidence in**  
**Judgments of Ambiguity**  
**(N = 104)**

<u>Source of Variation</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>	<u>F</u>
Mode	1	.12	.12	< 1
Subjects (Mode)	104	652.61	6.28	
Type	1	12.85	12.85	16.89**
Mode X Type	1	.51	.51	< 1
Group	3	11.51	3.84	5.04**
Sentences (Group)	36	133.60	3.71	4.88**
Mode X Group	3	8.19	2.73	3.59*
Type X Group	3	14.55	4.85	6.38**
Mode X Type X Group	3	.36	.12	< 1
Error	4076	3100.66	.76	

\*  
p < .05  
\*\*  
p < .01

four groups (surface structure 54.40% correct, 6.31 mean confidence; deletion 52.22% correct, 6.34 mean confidence; deep structure 58.66% correct, 6.36 mean confidence; case 47.03% correct, 6.45 mean confidence).

Although the differences in judgments and confidence ratings for the visual and auditory modes of presentation did not differ significantly, the mode X type interaction for judgments was significant ( $F = 8.39, p < .01$ ). From Figure 5 where this interaction is displayed it can be seen that the probability for the realization of the ambiguity potential of sentences was higher when they were presented in the visual than when in the auditory mode, whereas the accuracy of the perception of the unambiguous sentences was higher for the auditory than the visual mode. The reliability of this interaction has to be viewed with some suspicion until it can be confirmed in other studies. For the present, its effects must be limited to the set of sentences examined in the experiment. So in summary, subjects' perceptions of syntactic ambiguity are a function of the source of that ambiguity, and the mode of presentation of the stimuli has no significant effect upon those perceptions.

The fact that the unambiguous sentences are more likely to be perceived as being unambiguous than ambiguous sentences are likely to be perceived as being ambiguous, regardless of mode of presenta-

tion, is anticipated by the simple truth that the ambiguity of some sentences is difficult to perceive, because one interpretation is so overwhelmingly probable that the existence of another structural description rarely occurs to the perceiver. When the results shown in Figure 5 are restated in terms of the percent of the subjects judging the unambiguous sentences to be ambiguous or meaningless (38.54% for the auditory mode and 40.85% for the visual mode), it becomes obvious that the visual mode of presentation had the effect of increasing the relative proportion of ambiguity judgments for both sentence types.

That the effect of sentences within groups is also significant for both the judgments ( $\underline{F} = 7.59, p < .01$ ) and the confidence ratings ( $\underline{F} = 4.88, p < .01$ ) is to be expected since no attempt was made to control for semantic or syntactic variation across sentences within groups.

Now we come to what might be considered the heart of the first experiment, the interaction of sentence types and groups for the ambiguity judgments ( $\underline{F} = 51.34, p < .01$ ) and for the confidence ratings ( $\underline{F} = 6.38, p < .01$ ). The means are recorded in Table 4 where it can be observed that the rank order of groups in terms of perception of ambiguity is that predicted by hypothesis (1), i.e. surface structure > deletion > deep structure > case relationships, though not all the

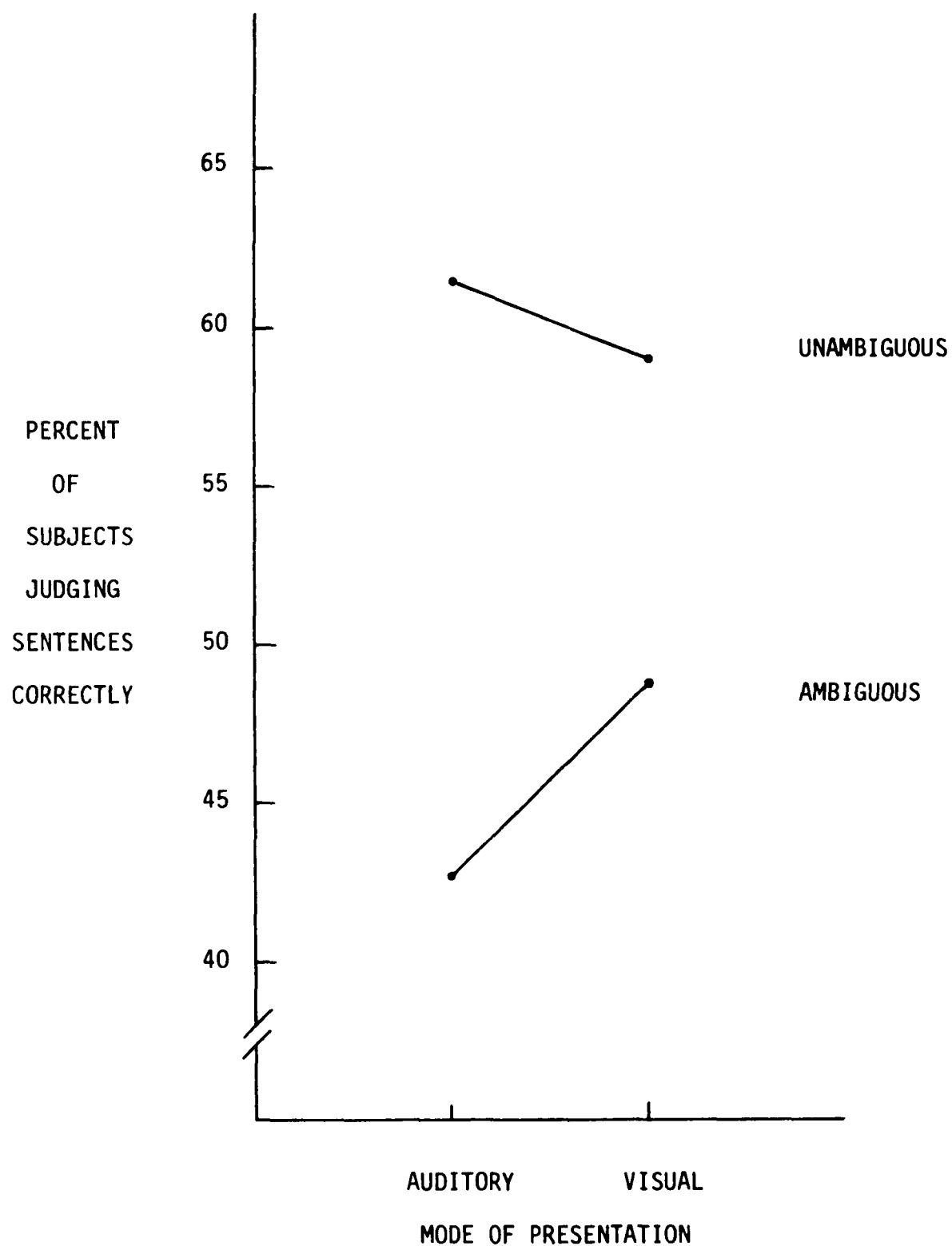


Figure 5. Mode X Type Interaction Effect on Judgments of Ambiguity



**TABLE 4**  
**Percent of Subjects Judging Sentences Ambiguous**  
**and Mean Confidence Ratings<sup>1</sup>**

<u>Sentence Group</u>	<u>Sentence Type</u>	
	<u>Ambiguous</u>	<u>Unambiguous</u>
Surface Structure	53.80%	45.01%
	6.34	6.27
Deletion	53.50%	49.06%
	6.31	6.37
Deep Structure	50.47%	33.14%
	6.48	6.24
Case Relationships	25.33%	31.32%
	6.54	6.35

<sup>1</sup>Figures for the unambiguous type sentences were obtained by subtracting from 100% the percent of judgments that were "correct" so consequently the percentages shown include judgments that the sentences were ambiguous and judgments that the sentences were meaningless.

differences are statistically significant, the least significant difference being 4.03%. In view of the fact that the ambiguity deriving from the surface structure and deletion sources can in a sense be resolved at the level of surface structure, the similarity between the means of these two groups is not surprising.

Of the ten sentences ambiguous in surface structure relationships, that most frequently perceived as such was

(23B) The foolish boy gave her cat food (75.47%)

and that most infrequently perceived as ambiguous was

(7B) The pretty girl's coat was sent back (31.48%).

As for sentences whose ambiguity arises from deletion transformations, that most frequently perceived was

(39A) John knows a taller man than David (80.77%)

and that least often perceived as ambiguous was

(12B) He sold his modern paintings and furnishings (18.52%).

That the ambiguity of deep structure relationships was less frequently perceived attests to their greater subtlety, an observation which is compatible with MacKay and Bever's finding (1967) that subjects needed more time to detect deep structure than surface structure ambiguities. In the present experiment the ambiguity of

(53B) The Spanish scholar wrote a popular book

and

(46A) The French teacher enjoys sports and music were seldom perceived, 14.81% and 26.92% respectively. Not coincidentally, these sentences and their unambiguous counterparts were among those eliminated in conducting the second experiment. The deep structure ambiguity of

(37A) These friends are the ones to help was, however, often realized (by 76.92% of the subjects).

Tabulation of the ambiguity judgments also shows that the sentences which were constructed to be ambiguous in case relationships were infrequently perceived as being ambiguous by the experimental subjects. At one extreme is

(35A) The green acid quickly dissolved the metal<sup>2</sup> which was judged to be ambiguous only 9.62% of the times it was presented. The case sentence whose ambiguity potential was most likely to be realized was

(48A) Their big strong brother hit the windshield (61.54%). In this instance it seems that one could argue, with considerable justification, that the ambiguity of the sentence originates in the lexical

---

<sup>2</sup>According to Langendoen (1969:27) the two interpretations can be paraphrased as (35A') The metal dissolved the green acid quickly and (35A'') [The animate agent] quickly dissolved the metal with the green acid. In other words, the question revolves around whether the relationship between acid and dissolved is locative or instrumental.

interpretation assigned to hit, i. e.  $\pm$  purposive, and it is this choice which determines the case relationships, not the case relationships which determine the meaning.<sup>3</sup>

Several interesting things become apparent when one looks at the group means for the unambiguous sentences presented in Table 4. These controls were judged in accordance with expectations, i. e. less ambiguous than their counterparts, for the surface structure and deep structure groups, and for the deletion group too though the difference between the ambiguous and unambiguous sentences in this category is not very large. Nevertheless in each of these three groups at least one of the unambiguous sentences was judged to be ambiguous more frequently than its counterpart. In the ten pairs of case structure sentences this reversal occurred six times.<sup>4</sup> The fact that subjects were so persistent in finding more than one interpretation for sentences reinforces the previously mentioned notion that if one works at it, more than one meaning can be "discovered" for almost any sentence. Therefore if further research is conducted with a similar paradigm, additional precautions should be taken in constructing the "unambiguous" control sentences to reduce the chances that this might

---

<sup>3</sup>For further explication of this viewpoint refer to Lakoff (1968 and Kooij (1971:67ff).

<sup>4</sup>The item numbers of these sentences are: 7, 31; 13, 51, 6, 12; 53; 11, 16, 49, 15, 22, 21.

happen.

Since there may be some interest on the part of researchers in using ambiguous sentences for which perceptual norms have been established, the percent of subjects ( $N = 106$ ) judging each of the present sentences ambiguous is shown in Appendix A. One way to combine the information from the ambiguity judgments and the degree of confidence expressed in them is to compute, for each sentence, an index of ambiguity which is similar to the index of comprehensibility proposed by Schwartz, Sparkman and Deese (1970). This involves making the calculations expressed in the following formula:

$$\text{index of ambiguity} = A (MC_A) - U (MC_U)$$

where  $A$  is the relative proportion of judgments that a sentence is ambiguous expressed in decimal form,  $MC_A$  is the mean confidence rating for those judgments,  $U$  is the relative proportion of judgments that the sentence is unambiguous, and  $MC_U$  is the mean confidence expressed in the unambiguous judgments. Indices such as this are also listed in Appendix A with the test sentences.

Turning now to the confidence ratings shown in Table 4, it is interesting to note that the highest means were reached for sentences ambiguous in deep structure and case relationships, the two groups perceived most inaccurately. So quite contrary to what one might expect, the subjects were most sure of themselves when they were wrong.

Common sense would also lead one to believe that the mean confidence ratings would be higher for ambiguous than for unambiguous sentences, since once a subject detects a second interpretation of a sentence he could be expected to demonstrate great confidence in his judgment of its ambiguity. This was indeed the situation for the surface structure, deep structure, and case relationship sentences, but not for the deletion group.

No explanation is offered for the mode X group interaction effect upon the confidence ratings though it can be mentioned that it is barely significant at the .05 confidence level and that the degree of confidence was higher for the auditory mode of presentation of sentences in the deletion and case relationship groups, whereas the confidence expressed in judgments as to the ambiguity of sentences in the surface structure and deep structure groups was greater when these sentences were presented visually.

### Comprehensibility

If it is true that the processing of ambiguous sentences includes the conscious or unconscious assignment of more than one structural description to a given sequence, then there exists the possibility that this intricacy might be reflected in lower comprehensibility ratings for ambiguous than for unambiguous sentences. A problem arose, however, because some subjects (N = 38) in this task set made only

the obvious distinction between the ten anomalous sentences and the others; they classified the former as incomprehensible and expressed the highest degree of confidence in those judgments. Then the remaining forty sentences were uniformly categorized as comprehensible with the highest confidence. Consequently the results of this portion of the first experiment do not reveal differences of the same magnitude as those just presented. Nevertheless analyses of variance were computed and the results of these statistical tests are presented in Tables 5 and 6.

The means for the two sentence types are such that the ambiguous sentences were judged comprehensible slightly less frequently than the unambiguous sentences, by 93.76% and 94.35% of the subjects, respectively, but the difference is not statistically significant. However, it should also be noted that of the 40 pairs of sentences under consideration, in 23 of these pairs the unambiguous member was judged comprehensible more frequently than the ambiguous member, in 11 pairs the reverse was true, and the percentages were identical for 6 of the pairs.<sup>5</sup> According to the tables then, neither the mode of presenta-

---

<sup>5</sup>The item numbers are: 6, 9, 10, 13, 15, 18, 20, 21, 23, 27, 29, 30, 32, 33, 34, 37, 41, 42, 43, 44, 47, 51, 54; 7, 12, 16, 19, 22, 31, 36, 39, 49, 50, 55; and 11, 25, 35, 46, 48, 53, respectively. It should be kept in mind that the sizes of these differences within pairs, often quite small, are disregarded for the moment and only the direction is considered.

**TABLE 5**  
**Analysis of Variance for Comprehensibility Judgment**  
**(N = 102)**

<u>Source of Variation</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>	<u>F</u>
Mode	1	.03	.03	< 1
Subjects (Mode)	100	34.02	.34	
Type	1	.04	.04	< 1
Mode X Type	1	.22	.22	4.94 <sup>*</sup>
Group	3	4.44	1.48	33.60 <sup>**</sup>
Sentence (Group)	36	14.69	.41	9.25 <sup>**</sup>
Mode X Group	3	.48	.16	3.63 <sup>*</sup>
Type X Group	3	.33	.11	2.53
Mode X Type X Group	3	.68	.23	5.12 <sup>**</sup>
Error	3916	172.67	.04	

<sup>\*</sup>p < .05  
<sup>\*\*</sup>p < .01



**TABLE 6**  
**Analysis of Variance for Degree of Confidence in**  
**Comprehensibility Judgments**  
**(N = 102)**

<u>Source of Variation</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>	<u>F</u>
Mode	1	26.45	26.45	2.83
Subjects (Mode)	100	935.91	9.36	
Type	1	.35	.35	< 1
Mode X Type	1	1.12	1.12	1.53
Group	3	54.57	18.19	25.01**
Sentences (Group)	36	232.18	6.45	8.87**
Mode X Group	3	2.51	.84	1.15
Type X Group	3	8.38	2.79	3.84**
Mode X Type X Group	3	7.21	2.40	3.31*
Error	3916	2848.06	.73	

\*  
 \*\*  
 p < .05  
 p < .01

tion nor the type of sentence has a significant effect upon the judgments of comprehensibility or the degree of confidence. Yet the F value for the interaction effect of these variables on the nature of the judgment is significant at the .05 confidence level, an effect which is clearly demonstrated in Figure 6.

Thus the hypothesis that ambiguous sentences tend to be judged less comprehensible than their unambiguous counterparts holds, if at all, only for the auditory mode of presentation. In the visual mode the ambiguous sentences were in fact judged comprehensible by a higher percentage of subjects than the unambiguous sentences. Because the magnitude of the differences is so small even in comparison with those displayed in Figure 5 for ambiguity judgments, no inferences can be drawn from this finding until additional evidence is obtained.

The differences among sentences within groups and the group means themselves were significant sources of variation in both analyses. The rank order of groups from most to least comprehensible is case relationships > deep structure > surface structure > deletion, and the same order prevailed for the confidence ratings which suggests that confidence in judgments of incomprehensibility tend to be lower than confidence in judgments that a sentence is comprehensible. As for the effect of the mode X group interaction on judgments (F = 3.63,  $p < .05$ ), suffice it to note that the surface structure and case sentences

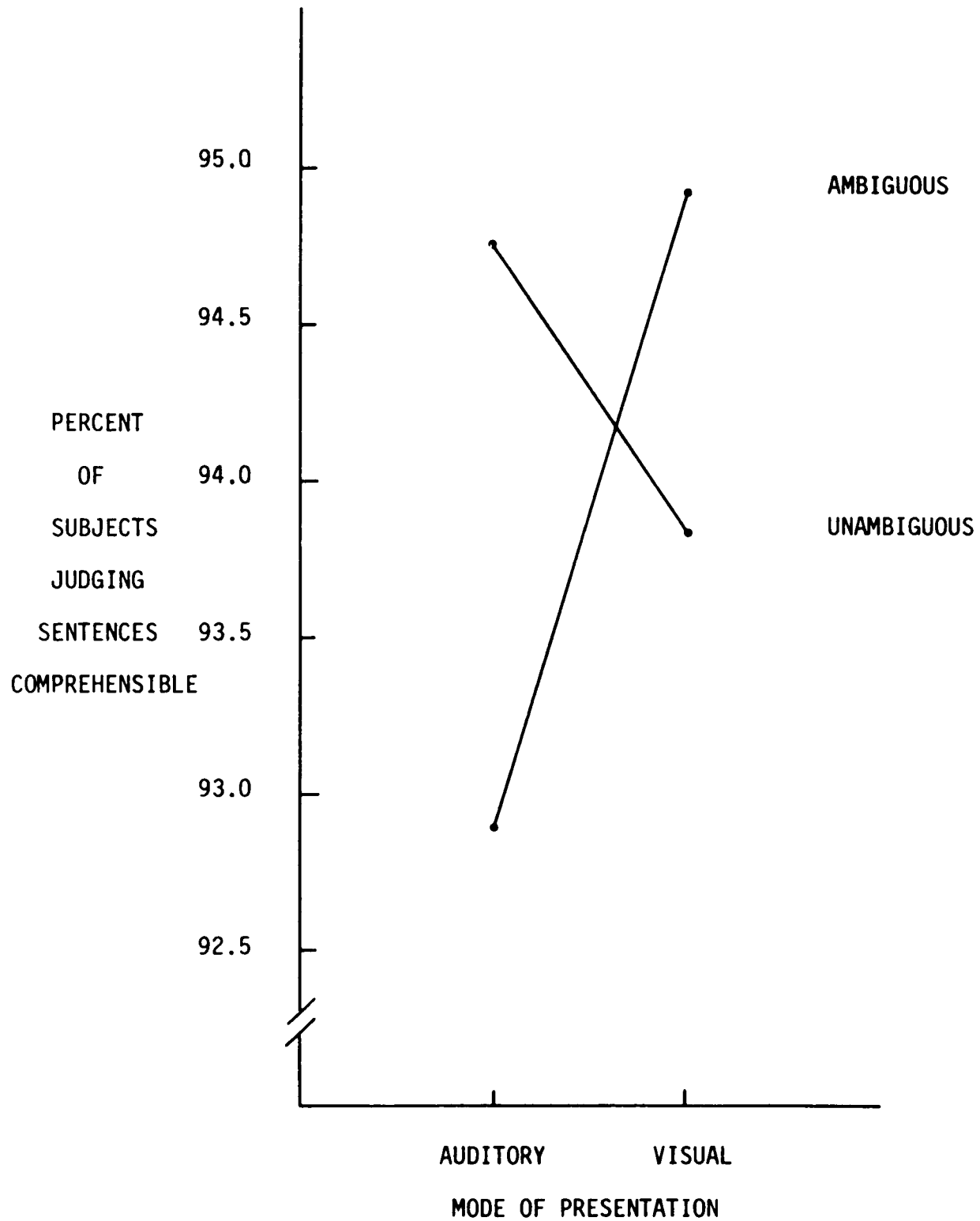


Figure 6. Mode X Type Interaction Effect on Judgments of Comprehensibility

were judged to be more comprehensible in the auditory presentation, and the deletion and deep structure sentences were judged more comprehensible when presented visually.

Looking now at the type X group interaction we see that although the  $\underline{F}$  value of 2.53 does not reach the desired confidence level, the direction of the differences for three of the four groups is that predicted by hypothesis (3). To be more specific, the unambiguous sentences of the surface structure, deletion, and deep structure groups were judged more comprehensible on the average than their respective ambiguous counterparts. That the ambiguity potential of the case structure sentences is so seldom realized may explain why the comprehensibility ratings of these sentences did not follow the same pattern. Here again it was observed that the higher the judged comprehensibility of a group of sentences, the higher is the mean confidence rating for that group. The extent of this relationship is indicated by  $\underline{r}_s = .95$  ( $p < .01$ , one-tailed test), the value of Spearman's rank correlation coefficient calculated for these variables. Another way of stating the relationship is that the higher the proportion of negative judgments (e.g. incomprehensible as opposed to comprehensible), the lower the confidence ratings. Furthermore, in view of the significance of the interactions we have just discussed, the effectiveness of the mode X type X group interaction is to be expected.

When one inquires as to which of the ambiguous sentences are judged highest in comprehensibility (by 100% of the subjects), it can be seen that these seven ambiguous sentences<sup>6</sup> are among the thirteen sentences with the lowest probabilities for realization of their ambiguity potential, all being under 34%. Also, for the five ambiguous sentences judged lowest in comprehensibility,<sup>7</sup> the probabilities for realization of the ambiguity potential were all greater than 34%, and three of the five exceeded 50%. But it was not found to be the case that the sentences whose ambiguity was most frequently perceived were judged lowest in comprehensibility. So it seems that ambiguity does have at least a weak effect on judgments of the comprehensibility of a sentence, but other properties may confound or override this effect.

### Complexity

Much of the evidence presented in Chapter II seems to indicate that the cognitive complexity of ambiguous sentences is greater than that of unambiguous sentences. Whether this difference can be reflected in judgments of the complexity of sentences was examined in this third and final phase of the first experiment. The behavior of some of the subjects (N = 10) assigned to this task was similar to that

---

<sup>6</sup>Items 11A, 12B, 31B, 35A, 46A, 49B, 50A.

<sup>7</sup>Items 6B, 13A, 23B, 44B, 51B.

previously described insofar as the uniformity of the ratings of non-anomalous sentences is concerned. Nevertheless some interesting findings are revealed in Tables 7 and 8.

Note first of all the main effect of sentence type on the judgment and degree of confidence ( $\underline{F} = 7.74$  and  $8.14$ , respectively,  $p < .01$ ). The mean scores for all presentations of the sentences indicate that the ambiguous stimuli were judged to be complex more frequently than were the unambiguous stimuli,  $20.26\%$  and  $17.02\%$  respectively. Furthermore, the mean confidence rating was higher for the ambiguous ( $6.47$ ) than for the unambiguous ( $6.38$ ) sentences. This is in accordance with the pattern of the previously reported finding with respect to confidence in comprehensibility judgments, inasmuch as it appears that the higher the proportion of negative judgments expressed, the lower the mean confidence ratings. So there is a tendency for greater confidence to be expressed in positive than in negative judgments.

The effects of mode, the mode X type, and mode X group interactions were not significant, thus providing substantial evidence for rejecting the third clause of hypothesis (2) and accepting a null hypothesis to the effect that the mode of presentation has no significant effect upon judgments of the complexity of sentences. Once again though, the differences among groups ( $\underline{F} = 37.92$  and  $21.71$ ) and among sentences within groups ( $\underline{F} = 9.40$  and  $8.44$ ) could not very probably be

**TABLE 7**  
**Analysis of Variance for Complexity Judgment**

(N = 82)

<u>Source of Variation</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>	<u>F</u>
Mode	1	.96	.96	< 1
Subjects (Mode)	80	93.74	1.17	
Type	1	.85	.85	7.74**
Mode X Type	1	.12	.12	1.12
Group	3	12.56	4.19	37.92**
Sentences (Group)	36	37.37	1.04	9.40**
Mode X Group	3	.38	.13	1.16
Type X Group	3	2.74	.91	8.28**
Mode X Type X Group	3	.37	.12	1.12
Error	3128	345.50	.11	

\*\* p < .01

TABLE 8  
 Analysis of Variance for Degree of Confidence in  
 Judgments of Complexity  
 (N = 82)

<u>Source of Variation</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>	<u>F</u>
Mode	1	4.86	4.86	< 1
Subjects (Mode)	80	874.13	10.93	
Type	1	6.19	6.19	8.14**
Mode X Type	1	.49	.49	< 1
Group	3	49.49	16.50	21.71**
Sentences (Group)	36	230.82	6.41	8.44**
Mode X Group	3	5.11	1.70	2.24
Type X Group	3	3.90	1.30	1.71
Mode X Type X Group	3	2.11	.70	< 1
Error	3128	2377.11	.76	

\*  
p < .05  
 \*\*  
 p < .01



attributed to chance (  $p < .01$ ).

The rank order of groups from most to least complex is deletion > surface structure > deep structure > case relationships. It should be pointed out that this is exactly the opposite of the ranking of groups in terms of comprehensibility. In other words, the group of sentences judged most complex is the same as that judged least comprehensible by a sample of subjects from the same population, that group judged least complex is most comprehensible, and so forth. This finding suggests, then, that in assessing the comprehensibility of a linguistic stimulus, one is also assessing its complexity and vice versa.

This assertion finds additional support in the effect of the type X group interaction on complexity judgments ( $F = 8.28$ ,  $p < .01$ ) because the pattern is the same as that reported for comprehensibility though the means did not differ significantly in that analysis. They did, however, in the present task set and the ambiguous sentences in the surface structure, deletion, and deep structure groups were judged complex more frequently than were the unambiguous sentences in those groups. The unambiguous case sentences were found to be less complex than their ambiguous counterparts, on the average, and, again, a plausible explanation for this reversal seems to be that the ambiguity of the case relationships is perceived so infrequently.

Now as attention is focused upon comparing the complexity ratings within sentence pairs, it becomes even more obvious that ambiguity does indeed affect the estimates of the complexity of a stimulus. In 20 pairs the ambiguous member was judged complex more frequently than the unambiguous member; in 14 pairs it was the unambiguous member that was judged more complex; and in six pairs the scores were tied or so close that the difference was negligible.<sup>8</sup> As was illustrated with the comprehensibility ratings, the ambiguous sentences which were consistently judged not complex (N = 4) were sentences with < 34% probability of being perceived as ambiguous. In fact, these four sentences were also among those highest in comprehensibility ratings.<sup>9</sup> And for 10 of the 12 ambiguous sentences most frequently judged complex, the probabilities for realization of the ambiguity potential were > 60%.<sup>10</sup>

In general then, ambiguity does have the predicted effect upon intuitive judgments of the complexity of sentences, provided that the

---

<sup>8</sup>The respective item numbers are: 9, 10, 20, 23, 29, 30, 32, 33, 34, 35, 37, 39, 41, 43, 44, 46, 48, 50, 54, 55; 6, 7, 12, 13, 16, 19, 21, 22, 31, 36, 42, 49, 51, 53; 11, 15, 18, 25, 27, 47.

<sup>9</sup>The four ambiguous sentences judged not complex by all the subjects are 21B, 31B, 49B, and 53B.

<sup>10</sup>These ten sentences are: 9A, 10A, 32A, 37A, 39A, 41A, 42A, 43A, 44B, 54A. The two remaining items are 13A and 51B.

ambiguity is not so intricate and subtle that it can only rarely be detected by native speakers. Thus the rating data provide support for the broad hypothesis that ambiguity contributes to the interpretational density of a linguistic stimulus. Though definitive conclusions cannot be reached on the basis of information available at this time, some results pertinent to developing a description of the actual processing of ambiguous sentences are presented in the remainder of this chapter.

Before leaving this discussion of the findings of the first experiment and turning to the second, it seems appropriate to summarize in terms of the experimental hypotheses. The data collected provide support for the following:

- (1) The recognition of ambiguous sentences as ambiguous is function of the syntactic source of the ambiguity with the following order from most to least readily recognized:  
surface structure > deletion > deep structure > case relationships.
- (4) Within each group of sentences the complexity rating for ambiguous type sentences is greater than the complexity rating for unambiguous type sentences. (This statement did not hold for the case relationship sentences.)
- (5) The sentences judged relatively low in comprehensibility

and relatively high in complexity are those with a relatively high probability for realization of the ambiguity potential.

On the basis of the present findings it is not possible, however, to support a stronger version of (5) to the effect that:

Sentences with a relatively high probability for realization of the ambiguity potential are judged relatively low in comprehensibility and high in complexity.

#### Hypotheses

- (2) The judgments of a sentence's ambiguity, comprehensibility, and complexity vary depending upon the mode of presentation (aural or visual) of the stimuli

and

- (3) Syntactic ambiguity affects native speakers' intuitions about sentences in such a way that within each group of sentences the ambiguous type sentences are judged lower in comprehensibility than their unambiguous counterparts

are rejected.

#### Experiment II

Regardless of the particulars of the human sentence processing model one favors, an essential component of the device is short-term memory. The objective of this experiment was to determine the effects of several varieties of syntactic ambiguity on short-term

memory capacity. The general hypothesis being tested asserts that one way the processing of ambiguous sentences differs from the processing of unambiguous sentences is that the former occupy more short-term memory space than the latter.

### Sentence Recall

First, however, we shall consider the results of an analysis of variance of the errors made in recalling the sentences. It will be remembered that because of time limitations and especially because of the contamination which would result if subjects learned both the ambiguous and the unambiguous member of each of the 32 pairs of sentences, two forms of the test were employed, one set of subjects taking Form A and another set taking Form B. Thus, differences between these two sets of subjects comprise a potential source of variation in the experimental results which must be examined. Table 9 shows that as far as sentence recall is concerned, the difference between sets is not statistically significant although set does interact with several other variables. The effect of the set X group interaction ( $F = 5.12, p < .01$ ) derives from the fact that one set of subjects performed better on one group of sentences, the other set performed better on another group of sentences, with no significant differences being revealed for the remaining two groups.

Sentence group and type are shown to be significant main effects

TABLE 9

## Analysis of Variance for Errors Made in Sentence Recall

(N = 71)

<u>Source of Variation</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>	<u>F</u>
Set	1	.04	.04	< 1
Subjects (Set)	69	9.63	.14	
Group	3	4.04	1.35	14.17**
Type	1	1.58	1.58	16.68**
Group X Type	3	.96	.32	3.37*
Set X Group	3	1.46	.49	5.12**
Set X Type	1	.35	.35	3.68
Set X Group X Type	3	5.04	1.68	17.69**
Sentence (Set-Group-Type)	48	51.27	1.07	11.24**
Error	2139	203.42	.10	

\*p < .05  
 \*\*p < .01

( $\underline{F} = 14.17$ ,  $\underline{F} = 16.68$ ,  $p < .01$ ) with the rank order of groups in terms of recall efficiency being deep structure (92.43% correct) > surface structure (85.74% correct) > case relationships (83.80% correct) > deletion (80.99% correct), the least significant difference being 2.53%. Contrary to the prediction made in hypothesis (1), the ambiguous sentences were correctly recalled more frequently (88.38%) than the unambiguous sentences (83.10%). This finding becomes even more pronounced when one looks at the group X type interaction ( $\underline{F} = 3.37$ ,  $p < .05$ ) which reveals that within the deletion, deep structure, and case relationship groups, the means for the recall of ambiguous sentences are significantly higher than the means for the unambiguous sentences. In the surface structure group, recall was slightly but not significantly better for the unambiguous sentences. Moreover, the differences cannot be attributed to subject sets since the interaction of set and sentence type is not significant ( $\underline{F} = 3.68$ ).

This interaction is further substantiated when one separates the group X type responses according to subject sets ( $\underline{F} = 17.69$ ,  $p < .01$ ) for then it can be seen that the subjects in set 2 recalled the ambiguous sentences better than the unambiguous sentences within each of the four groups, although the means in the deletion and deep structure groups do not differ significantly. Interestingly, the recall of sentences in these two groups by subjects in set 1 followed the same

pattern and in fact, the differences in this case are significant. But these subjects correctly recalled the unambiguous sentences in the surface structure and case relationship groups more frequently than they did the ambiguous sentences although in the latter category, the difference was not statistically significant. When one remembers that the sentences ambiguous in case relationships are not likely to be so perceived, the meaningfulness of the trend diminishes. Hence, the direction of six of the eight possible comparisons (sentence types within each of the four groups of sentences for both sets of subjects) is opposite to predictions and only one, the better recall of unambiguous than ambiguous sentences in the surface structure group by subjects in set 1, was predicted. Therefore, hypothesis (2) is rejected.

Since it is possible to argue that the differences between groups of sentences can be attributed to semantic rather than or in addition to syntactic variation, it appears productive to move to another level and examine the relationship between the scores of members of each pair of sentences. It will be recalled that every effort was made to use the same words in constructing the ambiguous and the unambiguous counterparts. Therefore, if the scores for these sentences differ significantly, at least some of this difference can be attributed to the ambiguity. Thus, although a hypothesis regarding within pair com-



parisons was not formulated, it is still interesting to note that in 19 pairs the recall of ambiguous sentences was superior to that of unambiguous sentences, in 7 pairs the unambiguous sentences were recalled more efficiently, and the scores did not differ significantly in 6 pairs.<sup>11</sup> That differences among test items were a significant source of variation ( $F = 11.24$ ,  $p < .01$ ) was anticipated and it would be astonishing if such differences had not occurred, given that data from the first experiment revealed substantial sentence-to-sentence variation in the ratings of ambiguity, comprehensibility, and complexity.

Before attempting to explain the patterns of recall errors, a few remarks concerning the nature of the errors are necessary. One of the most common non-serious errors was the transposition of formatives from the same grammatical class. For example,

(51A) Observers thought Sue won and Mary did too  
was frequently recalled as

Observers thought Mary won and Sue did too.

Since this sort of mistake did not reflect a change in the syntactic structure of a sentence nor its ambiguity potential, it was not considered a major error. It is merely mentioned here as an interesting

---

<sup>11</sup>The respective item numbers are: 7, 10, 11, 12, 15, 16, 18, 19, 20, 22, 27, 29, 32, 34, 36, 37, 39, 48, 50; 9, 25, 30, 43, 44, 51, 55; 23, 33, 42, 47, 49, 54.

phenomenon that appears worthy of further investigation.

Table 10 shows that the total number of serious errors made was 324, which works out to an average of 3.59 per subject included in the analysis. The incomplete sentences probably result largely from interference occurring during the listening process, inasmuch as when the tapes were played in the classroom extraneous noises such as the ringing of the quarter-hour chimes, coughing, and sneezing did occur. The same explanation can also be advanced to account for errors involving the omission of a single noun. In general, when subjects made this sort of mistake, they indicated their awareness of it by leaving a blank space in place of the missing word, for example

(34B) \_\_\_\_\_ likes money more than her husband.

However, when qualifiers were omitted, a frequent occurrence in a sentence such as

(11A) The very handsome young soldier was crippled

or

(11B) The very handsome young soldier was seized,

the subjects generally seemed unaware of the discrepancy between the stimulus and their response. Thus it appears to be the case that under conditions of memory overload priority is placed upon retaining the verb or verbs, then the nouns and/or pronouns, and lastly the qualifiers. Further evidence of the relatively greater uncertainty

TABLE 10  
Nature of Sentence Recall Errors  
(N = 71)

<u>Error</u>		<u>Number</u>	<u>Percent</u>
Incomplete Sentences		85	26
Omissions of a Single Word		96	30
Qualifier	68		
Noun	18		
Pronoun	7		
Verb	3		
Addition of a Word		9	3
Qualifier	6		
Noun	2		
Verb	1		
Change in Ambiguity Potential		89	27
Unambiguous to Ambiguous	35		
Ambiguous to Unambiguous	54		
Transformational Changes		7	2
Active to Passive	4		
Passive to Active	3		
Miscellaneous		<u>38</u>	<u>12</u>
TOTAL		324	100

about qualifiers results from the observation that items in this category were added to recalled sentences six of the nine times an error like this was made. Also, many of the errors which resulted in a modification of the ambiguity potential involved qualifiers. This was especially true of the sort of coordinate relationships expressed in sentences like

(50A) Their valuable recordings and films were stolen  
and

(50B) Their recordings and valuable films were stolen.

Why is it that ambiguity appears to enhance the recall of sentences? Might the fact that some of the ambiguous sentences are so obviously so, even to the point of being humorous, be responsible for the more efficient recall? Probably, this is not the explanation since this property could be attributed to so few of the test sentences. Perhaps the answer lies not in the source of the ambiguity potential nor the probability for realization of the potential but rather in linguistic properties of the sentences which affect recall so strongly as to nullify the effect of the ambiguity.

In order to ascertain whether sentence length might be one of these variables, the relationship between the number of alphabetic characters in a sentence and sentence recall was investigated. Although the F value of 15.22 is significant at the .05 confidence level,

the coefficient of correlation ( $r = -.0816$ ) is very low. Therefore, to a very limited, almost negligible extent, the longer the sentence is, the less likely it is to be correctly recalled. And since comparisons indicate that the unambiguous members are longer than the ambiguous counterparts in 19 pairs, as opposed to 9 pairs in which the reverse situation prevails, this tentative explanation based on length must be discarded.

Since meaningfulness of the stimulus is known to enhance the learning of verbal material, the ratings of sentence pairs along this dimension were examined. In nine pairs of sentences the unambiguous sentences were judged meaningless more frequently than their ambiguous counterparts. For eight of these nine pairs, recall of the ambiguous, apparently more meaningful, sentence was superior to recall of the unambiguous sentence.

In spite of the fact that every effort was made to obtain equally clear and interpretable recordings of the sentences, interpretation of the error data is complicated by the possibility that this goal may not have been achieved. For this reason enumeration and discussion of other factors which may have influenced the experimental results will be postponed until the digit recall data are reviewed.

### Digit Recall

The mean number of digit pairs recalled was 3.167 and the two

sets of subjects did not differ significantly from each other in their performance, as can be observed in the analysis of variance results presented in Table 11. However, the effect of sentence groups was significant ( $\underline{F} = 12.86$ ,  $p < .01$ ) with an average of 3.493 pairs being recalled after deep structure sentences, 3.139 after deletion sentences, 3.006 after surface structure sentences, and 3.000 after sentences in the case relationship group, the least significant difference being .1313. Stated differently, this means that the rank order of sentence groups in terms of heaviest to lightest short-term memory load is case relationship > surface structure > deletion > deep structure. This is rather surprising in view of the fact that the case sentences were rated the most comprehensible and the least complex of the four groups.

Note too that the effects of sentence type and the group X type interaction were not significant sources of variation. Therefore, although the differences in digit recall between sentence types on an overall basis and within the surface structure, deep structure, and case relationship groups are in the predicted direction, they are too small to stand the test of significance and all clauses of hypothesis (3) are rejected.

The set X type interaction ( $\underline{F} = 7.34$ ,  $p < .01$ ) which is portrayed in Figure 7 is rather interesting for it seems that subjects in Set 2 were indeed able to recall more digit pairs after the unambiguous than

TABLE II  
Analysis of Variance for Digit Recall  
(N = 71)

<u>Source of Variation</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>	<u>F</u>
Set	1	1.19	1.19	< 1
Subjects (Set)	69	1758.14	25.48	
Group	3	82.14	27.38	12.86**
Type	1	.03	.03	< 1
Group X Type	3	5.39	1.80	< 1
Set X Group	3	8.38	2.79	1.31
Set X Type	1	15.64	15.64	7.34**
Set X Group X Type	3	75.64	25.21	11.84**
Sentence (Set-Group-Type)	48	350.65	7.31	3.43**
Error	1815	3864.25	2.13	

\*\*  $p < .01$

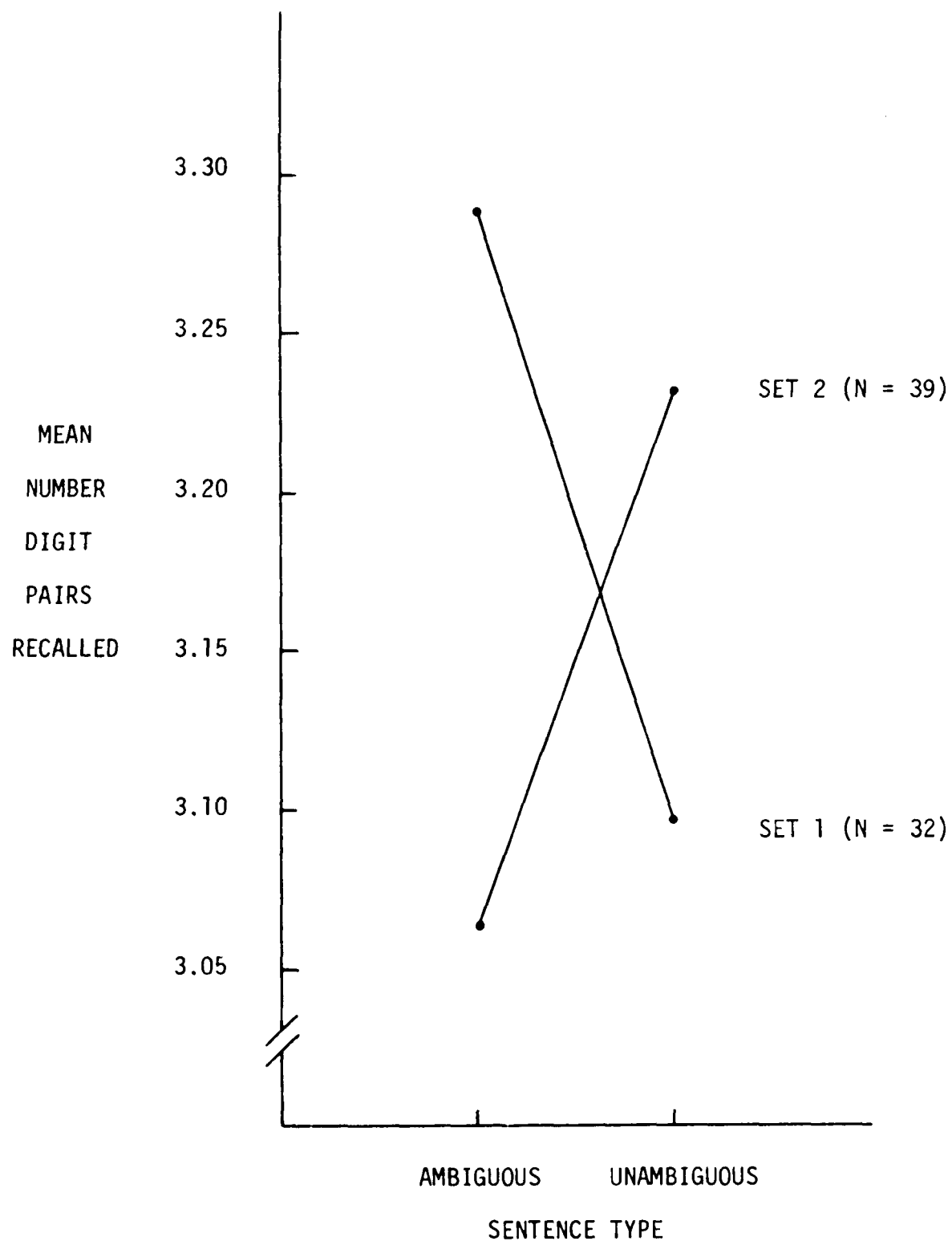


Figure 7. Set X Type Interaction Effect on Digit Recall



after ambiguous sentences whereas subjects in Set 1 recalled more digits after the ambiguous stimuli. When the means are broken down even further, i.e. types within groups by sets ( $F = 11.84$ ,  $p < .01$ ) it was observed that the behavior of the subject sets was not alike for any of the four groups. In other words, for subjects in Set 2, digit recall was higher after the unambiguous sentences in the deletion and deep structure groups, but the types reversed themselves for Set 1 subjects with the same groups of sentences. On the other hand, Set 2 subjects recalled more digit pairs after ambiguous sentences in the surface structure and case groups, whereas Set 1 subjects performed better in recalling digits after the unambiguous sentences in these groups. Whether this finding reflects real differences in the performance of the two groups of subjects or an inequality of the difficulty of parallel sections of Forms A and B of the test remains obscure.

The next step is to compare the recall of digits after the ambiguous and unambiguous counterpart within each pair. Not only were there relatively few significant differences in the 32 pairs, but also the stability of the means varied widely because the number of digit pairs recalled was scored only if the sentence preceding the digits was recalled without serious error. Nevertheless, here are the results of the tally: in 6 pairs the mean number of digits recalled after the

unambiguous member was higher than the mean after the ambiguous member, in 8 pairs digit recall was superior after the ambiguous member, and the scores did not differ significantly in 18 pairs.<sup>12</sup> In five of those six instances where the ambiguous sentences seemed to occupy more short-term memory space than their unambiguous counterparts, the probability for realization of the ambiguity potential of the sentences exceeded 50%.<sup>13</sup> However, it does not follow that sentences with a high probability for being perceived as ambiguous are necessarily those which impose the greatest burden on short-term memory capacity.

The possibility that an individual's own perception of ambiguity is a determining factor in the short-term memory load imposed by a sentence must also be considered. In order to explore this notion, the data were analyzed with perceived ambiguity as the independent variable and pairs of decimal digits recalled as the dependent variable. One flaw in this procedure needs to be kept in mind. It is that some subjects perceived very few sentences as being ambiguous, so their mean scores are based on a small number of trials. In order to compensate in some measure for this situation, data from subjects (N = 14)

---

<sup>12</sup>The item numbers are 9, 15, 23, 34, 48, 55; 10, 16, 19, 22, 43, 44, 49, 51; 7, 11, 12, 18, 20, 25, 27, 29, 30, 32, 33, 36, 37, 39, 42, 47, 50, and 54, respectively.

<sup>13</sup>These items are 9, 23, 34, 48, and 55.

who perceived the ambiguity of less than 5 of the 16 ambiguous sentences were disregarded in making the following comparisons. Ties between digit recall scores after ambiguous and unambiguous sentences occurred for four of the remaining 57 subjects, leaving data from 53 for analysis in a sign test. Calculations based upon 25 changes in the predicted direction, i. e. the number of digits recalled after unambiguous sentences being higher than the number of digits recalled after ambiguous sentences, and 28 in the opposite direction yielded a z value of  $-.028$  and resulted in the rejection of experimental hypothesis (5).

The number of alphabetic characters in a sentence provides a crude approximation of the delay between the presentation and recall of the digits. In other words, the more time was needed to transcribe a sentence, the longer the digits had to be stored. Furthermore, if decay from memory is a simple function of retention time, then we would expect to find an inverse relationship between the number of characters in a sentence and the number of pairs of decimal digits correctly recalled. In fact, however, the coefficient of correlation has a positive value,  $\underline{r} = .0696$ . Although the F value ( $\underline{F} = 9.479$ ) is significant at the .01 confidence level, the proportion of the total variance which can be attributed to this source is quite small.

In comparing these results with some earlier findings based

upon word recall, a somewhat questionable comparison for reasons soon to be considered, it appears that in the present situation as in the studies reported by Matthews (1968), forgetting is not simply a function of time. In fact, both experiments provide data contradictory to Glucksberg and Danks' (1969) contention that word recall is a function of word latency. Some obscurity does arise, however, in light of the fact that sentence length and the time elapsing between the presentation and the recall of the items from the list are confounded in all of these experiments, and thus it is impossible to separate the effects of one from the other.

One reason why it is difficult to compare the findings from experiments based on word recall with the present findings based upon digit recall is that different behavioral strategies may have been employed with the two types of stimuli. Nevertheless the procedures were very similar, inasmuch as the order of word categories and the order of the first member of the digit pairs were constant for all trials. In either case subjects could attend more closely to items at the beginning or the end of the lists. A third alternative open only to the subjects working with the digits consisted of attending to only the second member of each pair of digits and grouping these unpredictable digits into larger units, for instance as a telephone number, a technique which would result in enhanced recall. This strategy could not

be used with words since the recall of each was necessarily an all-or-none proposition. An indication of whether or not a subject used the third strategy was provided by the sequence of the recalled digits. If they were written in the same order as they were presented, it is likely that this technique was used. Otherwise, the order of recall generally reflected either recency or primacy effects.

Summing up then, the following null hypotheses are accepted and all six of the research hypotheses are rejected:

- (1') Unambiguous sentences are recalled no more frequently than ambiguous sentences.
- (2') Within each group of sentences, the number of errors in the recall of ambiguous sentences is no greater than the number of errors in the recall of unambiguous sentences.
- (3') Within each group of sentences (across all subjects), the mean number of digits recalled after the unambiguous sentences is no greater than the mean number of digits recalled after the ambiguous sentences.
- (4') Across all four groups of sentences and across all subjects (i. e. within each pair of sentences), the mean number of digits recalled after the unambiguous sentences is no greater than the mean number of digits recalled after the ambiguous sentences.

- (5') Across all four groups of sentences and within each subject, the mean number of digits recalled after unambiguous sentences is no greater than the mean number of digits recalled after sentences perceived as being ambiguous.
- (6') The time which elapses between the presentation and recall of digits as measured by the number of alphabetic characters transcribed per sentence is not inversely related to digit recall.

So in a memory overload situation like that set up in Experiment II where the emphasis is on verbatim recall and the processing time is limited, the ambiguous sentences seem to have been processed as if they were unambiguous. This evidence supports Foss, Bever, and Silver's conclusion (1968) that only one interpretation is generally assigned to an ambiguous sentence. However, in a different situation, for instance like that in Experiment I where more time was available for responding to the stimulus and the task consisted of judging rather than remembering, the ambiguity effectively increased the rated complexity.

One reason that none of the predicted differences were revealed in the second experiment may be that in a short-term recall task the subjects attempt to store surface structure strings without assigning structural descriptions to them. Also, experimental conditions such

as the order of presentation of the sentences, the time interval between the presentation and recall, and the relative order in which the sentences and the list are recalled, i.e. list or sentence first, should be more closely examined as potential sources of variation. This can be rather easily accomplished by slightly modifying the design of future experiments. Control of linguistic variables is much more difficult, but besides or in addition to ambiguity and sentence length there are several which may have affected sentence recall and the short-term memory load. Some of these (with references to research showing their effects on sentence processing) are: transformational complexity (Savin and Perchonock, 1965), lexical complexity (Fodor and Garrett, 1967), Yngve depth (Martin and Roberts, 1966), lexical density (Perfetti, 1969), and subject-verb contiguity (Blumenthal, 1966). Each of these factors contributes to the interpretational density of a sentence, and by isolating and sorting out their effects on sentence recall and short-term memory capacity we can move toward the development and formulation of an accurate account of linguistic behavior.

## CHAPTER V. SUMMARY AND CONCLUSIONS



First we must be much more careful than in the recent past to study the nature of our intuitions about sentences. It will not do to multiply formal 'levels' or apocalyptically to attack each others' theories every time a new kind of intuition appears on the horizon. We must first make sure that we understand its nature and its interaction with other sets of intuitions. Not only will this solidify the factual basis of linguistic description, it may offer some understanding of how further to develop experimental phenomenology. --Thomas G. Bever, "The Integrated Study of Language Behaviour" in Biological and Social Factors in Psycholinguistics edited by John Morton, 1970, p. 203.

In Experiment I about 300 high school seniors judged either the ambiguity, the comprehensibility, or the complexity of half of 40 ambiguous-unambiguous pairs of sentences and indicated on a 7-point scale the degree of confidence they had in each judgment. Half of the subjects in each group heard tape recordings of the sentences and the other half read them as they were individually projected. Subjects in Experiment II participated in a short-term memory task which consisted of listening to a tape recording of some of the same sentences, each of which was followed by a string of seven pairs of digits and then recalling the sentence and as many numbers as possible.

Analysis of the results of these studies revealed that:

- (1) The probability that the ambiguity potential of a sentence will be realized is a function of the source of the ambiguity;
- (2) The mode of presentation, visual or aural, had no significant differential effect upon the probability that a sentence

would be recognized as ambiguous. This generalization holds true, of course, only for sentences used in the tests;<sup>1</sup>

- (3) The ambiguous sentences tended to be rated more complex than their unambiguous counterparts. This was particularly true of ambiguous sentences whose potential was likely to be realized;
- (4) The unambiguous sentences tended to be recalled no more efficiently than their ambiguous counterparts;
- (5) The ambiguous sentences tended to occupy about the same amount of short-term memory space as their unambiguous counterparts as measured by the recall of pairs of decimal digits following the sentences.

Given that only 40 ambiguous sentences from the entire universe of potentially ambiguous English sentences were subjected to testing, it seems unwise to generalize very far from these results. The methodological contribution of the present study is, however, more universal, and its significance lies in the demonstration of the fruitfulness of a design whereby the results of the direct elicitation of intuitive judgments are integrated with an indirect measure of the properties of the same stimuli.

---

<sup>1</sup>See Chapter I and some of the references cited there for examples of sentences with which mode does indeed make a difference.

Some investigators, e.g. Gamlin (mimeo, p. 22), have claimed that syntactic and semantic processing are separate. This may be true but the research reported here provides a reminder that the semantic properties of a sentence are in some measure dependent upon its syntactic properties. Or, if one adopts the generative semanticist position one could argue, I suppose, that the syntactic properties are functions of semantic readings. At any rate, then, a relationship of dependency exists and perhaps there is no better illustration of this than in the observation that syntactically ambiguous sentences are necessarily semantically ambiguous as well. Moreover, the results at hand reaffirm the necessity for linguistic and language processing theories, regardless of whether these be syntactically or semantically based, to deal with underlying structures as well as observable phenomena. Hence this research provides further support for the notion that sentences are processed on at least two levels--as linear sequences of superficial units and as representations of hierarchical relationships.

In reply to Garrett's question (1970), "does ambiguity complicate the perception of sentences?" one can reply in the affirmative insofar as this attribute is measured by means of rated complexity. As for the question of Carey, Mehler, and Bever (1970), "when do we compute all the interpretations of an ambiguous sentence?" one can

reply, not in a short-term memory task like that called for in Experiment II. But perhaps if more time elapsed between presentation of the sentence and presentation of the digits, the effect of the ambiguity would be revealed in the recall results. Too, additional experimentation might be conducted with a format more similar to that originally employed by Savin and Perchonock, i. e. verbal recall of sentences and word lists rather than written recall of sentences and digits. So since it is possible and even quite probable that the subjects attempted to learn the test sentences by rote and in so doing adopted an unnatural or uncommon strategy for treating them, the results are interpreted to reflect an artifact of the experimental procedure rather than a refutation of Bever, Lackner, and Kirk's (1969) model for linguistic processing whereby "the underlying structures of sentences are the primary units of immediate speech processing." This, then, seems to be a plausible explanation for the fact that the results of Experiment II did not conform to the general hypothesis that potentially ambiguous sentences by virtue of the fact that they bear more than one structural description, are more difficult to process than are unambiguous sentences. Consequently, the theoretical implications of these findings are somewhat restricted.

Furthermore, it would seem exceedingly presumptuous for this writer to recommend drastic revisions in either of the linguistic

theories used as a basis for formulating the test sentences on the grounds of negative results obtained in a single experiment. The implication of this position is not that grammars are unaccountable to performance data (Oller, Sales, and Harrington, 1970), but rather that the relationships between theories of linguistic competence and models of linguistic performance are indirect (Fodor and Garrett, 1966).

Between-subject variation in the perception of linguistic ambiguity appears to be a promising area for future investigation. Individual differences in age, sex, socioeconomic status, verbal fluency, vocabulary size, and authoritarianism might readily be examined in terms of potential relationships with ambiguity perception. Presently, research is underway to determine whether the perception of syntactic ambiguity by non-native speakers can be used as an indicator of language proficiency.

In the previous chapter the notion of interpretational density was introduced. Some of the linguistic variables deserving further consideration in light of their potential interaction with ambiguity or direct contribution to the complication of sentence processing are syntactic complexity, lexical density, lexical complexity, and subject-verb contiguity. The relationship between ambiguity and intelligibility is also interesting enough to warrant some study. And in all psycho-

linguistic work with linguistic ambiguity, the probability for realization of the ambiguity potential of each sentence should be considered as a major variable.

Not only should additional examples of the four sources of ambiguity investigated in this study be tested, but also other syntactic and semantic sources should be studied in terms of their effects on intuitive judgments and perceptual processes. And since ambiguity is a universal feature of human language, there is much to be gained by extending the scope of future studies to include languages other than English.

### Selected Bibliography

Arthur, Henry B., Jr. (1968). "Syntactic Ambiguity and Equivalence in Modern English," Dissertation Abstracts 28:3165A.

Bever, Thomas G. (1968). "Associations to Stimulus-Response Theories of Language," in Theodore R. Dixon and David L. Horton (eds.), Verbal Behavior and General Behavior Theory. Englewood Cliffs, New Jersey: Prentice-Hall. Pp. 478-494.

\_\_\_\_ (1970a). "The Influence of Speech Performance on Linguistic Structure," in Giovanni B. Flores d'Arcais and Willem J. M. Levelt (eds.), Advances in Psycholinguistics. Amsterdam: North-Holland. Pp. 4-30.

\_\_\_\_ (1970b). "The Integrated Study of Language Behavior," in John B. Morton (ed.), Biological and Social Factors in Psycholinguistics. Urbana: University of Illinois. Pp. 158-209.

\_\_\_\_, Jerry A. Fodor, and W. Weksel (1965). "Is Linguistics Empirical?" Psychological Review, LXII: 493-500.

\_\_\_\_, J. R. Lackner, and R. Kirk (1969). "The Underlying Structures of Sentences are the Primary Units of Immediate Speech Processing," Perception and Psychophysics 5:225-234.

Bloomfield, Leonard (1933). Language. New York: Holt, Rinehart, and Winston.

Blumenthal, A. L. (1966). "Observations with Self-Embedded Sentences," Psychonomic Science 6:453-454.

\_\_\_\_ (1967). "Prompted Recall of Sentences," Journal of Verbal Learning and Verbal Behavior 6:203-206.

Bolinger, Dwight (1971). "Contrast in Depth of Embedding," Journal of English Linguistics 5:29-30.

Carey, Peter W., Jacques Mehler, and Thomas G. Bever (1970). "Judging the Veracity of Ambiguous Sentences," Journal of Verbal Learning and Verbal Behavior 9:243-254. Reprinted as "When Do We Compute All the Interpretations of an Ambiguous Sentence?" in Flores d'Arcais and Levelt (eds.), Advances in Psycholinguistics. Pp. 61-75.

Chomsky, Noam (1957). Syntactic Structures. The Hague: Mouton. Janua Linguarum Series Minor No. 4.

\_\_\_\_\_ (1959). "Review of B. F. Skinner's Verbal Behavior," Language 35:26-58.

\_\_\_\_\_ (1961). "On the Notion 'Rule of Grammar,'" in R. Jakobson (ed.), Proceedings of the Twelfth Symposium in Applied Mathematics 12:6-24. Reprinted in J. Fodor and J. J. Katz (eds.), The Structure of Language. Englewood Cliffs, New Jersey: Prentice-Hall. Pp. 119-136.

\_\_\_\_\_ (1964). Current Issues in Linguistic Theory. The Hague: Mouton. Janua Linguarum Series Minor No. 38.

\_\_\_\_\_ (1966). Topics in the Theory of Generative Grammar. The Hague: Mouton. Janua Linguarum Series Minor No. 56.

\_\_\_\_\_ (1967). "The Formal Nature of Language," in Eric H. Lenneberg (ed.), Biological Foundations of Language. New York: John Wiley. Pp. 397-442.

\_\_\_\_\_ (1968). Language and Mind. New York: Harcourt, Brace & World.

\_\_\_\_\_ (1969a). "Deep Structure, Surface Structure, and Semantic Interpretation," Mimeo from Indiana University Linguistics Club.

\_\_\_\_\_ (1969b). "Some Empirical Issues in the Theory of Transformational Grammar," Mimeo from Indiana University Linguistics Club.

\_\_\_\_\_ (1970a). "Problems of Explanation in Linguistics," in Robert Borger and Frank Cioffi (eds.), Explanation in the Behavioral Sciences. Cambridge: Cambridge University. Pp. 425-451.

\_\_\_\_\_ (1970b). "Formal Discussion of 'The Development of Grammar in Child Language' by Wick Miller and Susan Ervin," in Mark Lester (ed.), Readings in Applied Transformational Grammar. New York: Holt, Rinehart and Winston. Pp. 41-50.

\_\_\_\_\_ and Morris Halle (1968). The Sound Pattern of English. New York: Harper & Row.



- Clark, H. H. and J. S. Begun (1968). "The Use of Syntax in Understanding Sentences," British Journal of Psychology 59:219-229.
- Coleman, E. B. (1964). "The Comprehensibility of Several Grammatical Transformations," Journal of Applied Psychology 48:186-190.
- \_\_\_\_\_ (1965). "Responses to a Scale of Grammaticalness," Journal of Verbal Learning and Verbal Behavior 4:521-527.
- Danks, Joseph H. (1968). "Some Factors Involved in the Comprehension of Deviant English Sentences," Dissertation Abstracts 29:08B:3100.
- \_\_\_\_\_ (1969). "Grammaticalness and Meaningfulness in the Comprehension of Sentences," Journal of Verbal Learning and Verbal Behavior 8:687-696.
- \_\_\_\_\_ and Sam Glucksberg (1970). "Psychological Scaling of Linguistic Properties," Language and Speech 13:118-140.
- Delack, John B. (1970). "On the Nature of Certain Fundamental Notions in Linguistic Theory," Papers in Linguistics 3:431-476.
- Dik, Simon C. (1968). Coordination: Its Implications for a Theory of General Linguistics. Amsterdam: North Holland.
- Dingwall, William Orr (1971). "On So-Called Anaphoric to and the Theory of Anaphora in General," Journal of English Linguistics 5:49-77.
- Downey, R. G. and D. T. Hakes (1968). "Some Psychological Effects of Violating Linguistic Rules," Journal of Verbal Learning and Verbal Behavior 7:158-161.
- Epstein, William (1967). "Some Conditions of the Influence of Syntactical Structure on Learning: Grammatical Transformation, Learning Instructions and 'Chunking,'" Journal of Verbal Learning and Verbal Behavior 6:415-419.
- \_\_\_\_\_ (1969). "Recall of Word Lists Following Learning of Sentences and of Anomalous and Random Strings," Journal of Verbal Learning and Verbal Behavior 8:20-25.

- Fillenbaum, Samuel (1970). "On the Use of Memorial Techniques to Assess Syntactic Structures," Psychological Bulletin 73:231-237.
- Fillmore, Charles J. (1968). "The Case for Case," in Emmon Bach and Robert Harms (eds.), Universals in Linguistic Theory. New York: Holt, Rinehart and Winston. Pp. 1-88.
- Flores d'Arcais, Giovanni B. and Willem J. M. Levelt (eds.) (1970). Advances in Psycholinguistics. Amsterdam: North Holland.
- Fodor, Jerry A. and M. Garrett (1966). "Some Reflections on Competence and Performance," in J. Lyons and R. J. Wales (eds.), Psycholinguistics Papers. Edinburgh: Edinburgh University. Pp. 135-154.
- \_\_\_\_\_ and \_\_\_\_\_ (1967). "Some Syntactic Determinants of Sentential Complexity," Perception and Psychophysics 2:289-296.
- \_\_\_\_\_, \_\_\_\_\_, and T. B. Bever (1968). "Some Syntactic Determinants of Sentential Complexity, II: Verb Structure," Perception and Psychophysics 3:453-461.
- \_\_\_\_\_ and J. J. Katz (eds.) (1964). The Structure of Language. Englewood Cliffs, New Jersey: Prentice-Hall.
- Forster, Kenneth I. and Leonie A. Ryder (1971). "Perceiving the Structure and Meaning of Sentences," Journal of Verbal Learning and Verbal Behavior 10:285-296.
- Foss, Donald J. (1969). "Decision Processes During Sentence Comprehension: Effects of Lexical Item Difficulty and Position upon Decision Times," Journal of Verbal Learning and Verbal Behavior 8:457-462.
- \_\_\_\_\_ (1970). "Some Effects of Ambiguity upon Sentence Comprehension," Journal of Verbal Learning and Verbal Behavior 9:699-706.
- \_\_\_\_\_, Thomas G. Bever, and M. Silver (1968). "The Comprehension and Verification of Ambiguous Sentences," Perception and Psychophysics 4:304-306.

- \_\_\_\_\_ and R. H. Lynch, Jr. (1969). "Decision Processes During Sentence Comprehension: Effects of Surface Structure on Decision Time," Perception and Psychophysics 5:145-148.
- Frank, Sheldon M. and Harry Osser (1970). "A Psycholinguistic Model of Syntactic Complexity," Language and Speech 13:38-53.
- Fraser, C., U. Bellugi, and R. Brown (1963). "Control of Grammar in Imitation, Comprehension and Production," Journal of Verbal Learning and Verbal Behavior 2:121-135.
- Gamlin, Peter J. (1971). "Sentence Processing as a Function of Syntax, Short Term Memory Capacity, the Meaningfulness of the Stimulus and Age," Language and Speech 14:115-134.
- \_\_\_\_\_ (mimeo). "Level of Listening Comprehension as a Function of the Two Process Variables: Syntax and Meaningfulness," to appear in Language and Speech.
- Garrett, Merrill F. (1970). "Does Ambiguity Complicate the Perception of Sentences?" in Flores d'Arcais and Levelt (eds.), Advances in Psycholinguistics. Pp. 48-60.
- \_\_\_\_\_ and J. A. Fodor (1968). "Psychological Theories and Linguistic Constructs," in R. R. Dixon and D. L. Horton (eds.), Verbal Behavior and General Behavior Theory. Englewood Cliffs, New Jersey: Prentice-Hall. Pp. 451-477.
- Gleason, H. A., Jr. (1965). Linguistics and English Grammar. New York: Holt, Rinehart and Winston.
- Gleitman, Lila R. and Henry Gleitman (1970). Phrase and Paraphrase: Some Innovative Uses of Language. New York: W. W. Norton.
- Glucksberg, Sam and Joseph H. Danks (1969). "Grammatical Structure and Recall: A Function of the Space in Immediate Memory or of Recall Delay?" Perception and Psychophysics 6:113-117.
- Hamilton, Helen W. and James Deese (1971). "Comprehensibility and Subject-Verb Relations in Complex Sentences," Journal of Verbal Learning and Verbal Behavior 10:163-170.
- Harman, G. (1967). "The Psychological Aspects of the Theory of Syntax," Journal of Philosophy 64:75-87.

Hockett, C. F. (1954). "Two Models of Grammatical Description,"  
Word 10:210-233.

\_\_\_\_\_ (1961). "Grammar for the Hearer," in Roman Jakobson (ed.),  
Structure of Language and its Mathematical Aspects. New York:  
American Mathematical Society. Pp. 220-236.

Honeck, Richard P. (1971). "A Study of Paraphrases," Journal of  
Verbal Learning and Verbal Behavior 10:367-381.

Householder, Fred W. (1972). "Review of Essentials of English  
Grammar by D. Terence Langendoen," Language 48:184-190.

Jacobs, Roderick and Peter S. Rosenbaum (eds.) (1970). Readings in  
English Transformational Grammar. Waltham, Mass.: Ginn-  
Blaisdell.

Jakobovits, Leon A. and Murray S. Miron (eds.) (1967). Readings in  
the Psychology of Language. Englewood Cliffs, New Jersey:  
Prentice-Hall.

Johnson, N. F. (1965). "Linguistic Models and Functional Units of  
Language Behavior," in S. Rosenberg, (ed.), Directions in  
Psycholinguistics. New York: Macmillan.

Johnson-Laird, P. M. (1969). "On Understanding Logically Complex  
Sentences," Quarterly Journal of Experimental Psychology  
21:1-13.

Kaplan, A. (1955). "An Experimental Study of Ambiguity and Context,"  
Mechanical Translation 2:39-46.

Katz, Jerrold J. and J. A. Fodor (1963). "The Structure of a Semantic  
Theory," in J. A. Fodor and J. J. Katz (eds), The Structure of  
Language. Pp. 479-518.

\_\_\_\_\_ and Paul M. Postal (1964). An Integrated Theory of  
Linguistic Descriptions. Cambridge, Mass.: M.I.T. Press.

✓  
Kucera, H. and W. N. Francis (1967). Computational Analysis of  
Present Day American English. Providence: Brown University.

Kooij, Jan G. (1971). Ambiguity in Natural Language. Amsterdam:  
North Holland.

- Lakoff, George (1968). "Instrumental Adverbs and the Concept of Deep Structure," Foundations of Language 4:4-29.
- \_\_\_\_\_. (1970). Irregularity in Syntax. New York: Holt, Rinehart and Winston.
- Langendoen, D. Terence (1969). The Study of Syntax. New York: Holt, Rinehart and Winston.
- \_\_\_\_\_. (1970). Essentials of English Grammar. New York: Holt, Rinehart and Winston.
- Lees, Robert B. (1960). "A Multiply Ambiguous Adjectival Construction in English," Language 36:207-221.
- Lenneberg, Eric (1967). The Biological Foundations of Language. New York: John Wiley.
- Levelt, Willem J. M. (1969). "The Scaling of Syntactic Relatedness: A New Method in Psycholinguistic Research," Psychonomic Science 17:351:352.
- \_\_\_\_\_. (1970). "A Scaling Approach to the Study of Syntactic Relations," in Flores d'Arcais and Levelt (eds.), Advances. Pp. 109-121.
- Levene, Harold F. "Children's Comprehension and Production of Ambiguous Sentences," Unpublished Dissertation, The Florida State University, 1970.
- Lyons, John and R. J. Wales (eds.) (1966). Psycholinguistics Papers. Edinburgh: Edinburgh University.
- MacKay, D. G. (1966). "To End Ambiguous Sentences," Perception and Psychophysics 1:426-436.
- \_\_\_\_\_. (1969). "Effects of Ambiguity on Stuttering: Towards a Model of Speech Production at the Semantic Level," Kybernetik 5:195-208.
- \_\_\_\_\_. (1970). "Mental Diplopia: Towards a Model of Speech Perception at the Semantic Level," in Flores d'Arcais and Levelt (eds.), Advances. Pp. 76-100.

- MacKay, D. G. and T. G. Bever (1967). "In Search of Ambiguity Perception and Psychophysics 2:193-200.
- MacKay, H. and M. D. Slesinger (1960). "Responses to Language: Judgments of Grammaticalness," International Journal of American Linguistics 26:275-282.
- Marks, L. E. (1967a). "Judgments of Grammaticalness of Some English Sentences and Semi-sentences," American Journal of Psychology 80:196-204.
- \_\_\_\_\_ (1967b). "Some Structural and Sequential Factors in the Processing of Sentences," Journal of Verbal Learning and Verbal Behavior 6:707-713.
- \_\_\_\_\_ (1968). "Scaling of Grammaticalness of Self-embedded English Sentences," Journal of Verbal Learning and Verbal Behavior 7:965-967.
- \_\_\_\_\_ and G. A. Miller (1964). "The Role of Semantic and Syntactic Constraints in the Memorization of English Sentences," Journal of Verbal Learning and Verbal Behavior 3:1-5.
- Marshall, J. C. (1965). "Syntactic Analysis as a Part of Understanding," Bulletin of the British Psychological Society 18:28-36.
- Martin, Edwin and Kelyn H. Roberts (1966). "Grammatical Factors in Sentence Retention," Journal of Verbal Learning and Verbal Behavior 5:211-218.
- Martin, James G. (1968). "Temporal Word Spacing and the Perception of Ordinary, Anomalous, and Scrambled Strings," Journal of Verbal Learning and Verbal Behavior 7:154-157.
- Matthews, W. A. (1968). "Transformational Complexity and Short Term Recall," Language and Speech 11:120-128.
- McCawley, James D. (1968). "The Role of Semantics in a Grammar," in E. Bach and R. T. Harms (eds.), Universals in Linguistic Theory. New York: Holt, Rinehart and Winston.
- \_\_\_\_\_ (1970). "Semantic Representation," in Paul L. Garvin (ed.), Cognition: A Multiple View. New York: Spartan. Pp. 227-247.

- McCawley, J. D. (1971). "Meaning and the Description of Languages," in Jay F. Rosenberg and Charles Travis (eds.), Readings in the Philosophy of Language. Englewood Cliffs, New Jersey: Prentice-Hall. Pp. 514-533.
- Mehler, Jacques and Thomas G. Bever (1968). "The Study of Competence in Cognitive Psychology," International Journal of Psychology 3:273-280.
- Mehler, Jacques and Peter Carey (1967). "Role of Surface and Base Structure in the Perception of Sentences," Journal of Verbal Learning and Verbal Behavior 6:335-338.
- \_\_\_\_\_ and \_\_\_\_\_ (1968). "The Interaction of Veracity and Syntax in the Processing of Sentences," Perception and Psychophysics 3:109-111.
- Miller, George A. (1956). "The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information," Psychological Review 63:81-97.
- \_\_\_\_\_ (1962). "Some Psychological Studies of Grammar," American Psychologist 17:748-762. Also in Jakobovits and Miron (eds.), Readings in the Psychology of Language. Pp. 201-218.
- \_\_\_\_\_ (1964a). "Language and Psychology," in E. H. Lenneberg, (ed.), New Directions in the Study of Language. Cambridge, Mass.: M.I.T. Press, Pp. 89-107.
- \_\_\_\_\_ (1964b). "The Psycholinguists," Encounter 23:29-37. Also in Charles E. Osgood and Thomas A. Sebeok (eds.), Psycholinguistics: A Survey of Theory and Research Problems. Bloomington: Indiana University Press, 1965. Pp. 293-307.
- \_\_\_\_\_ (1967). "Psycholinguistic Approaches to the Study of Communication," in David L. Arm (ed.), Journeys in Science: Small Steps--Great Strides. Albuquerque: University of New Mexico. Pp. 51-73.
- \_\_\_\_\_ and N. Chomsky (1963). "Finitary Models of Language Users," in R. D. Luce, R. R. Bush, E. Galanter (eds.), Handbook of Mathematical Psychology, Vol. II. New York: Wiley. Pp. 419-493.

- Miller, G. A., G. A. Heise, and W. Lichten (1951). "The Intelligibility of Speech as a Function of the Context of the Test Materials," Journal of Experimental Psychology 41:329-335.
- \_\_\_\_\_ and S. Isard (1963). "Some Perceptual Consequences of Linguistic Rules," Journal of Verbal Learning and Verbal Behavior 2:217-228. Also in Jakobovits and Miron (eds.), Readings in the Psychology of Language. Pp. 219-231.
- Morton, John (ed.) (1970). Biological and Social Factors in Psycholinguistics. Urbana: University of Illinois.
- Oller, John W., Jr., B. D. Sales, and R. V. Harrington (1970). "Toward Consistent Definitions of Some Psycholinguistic Terms," Linguistics 57:48-59.
- Perfetti, C. A. (1969). "Sentence Retention and Depth Hypothesis," Journal of Verbal Learning and Verbal Behavior 8:101-104.
- Postal, Paul M. (1964). "Underlying and Superficial Linguistic Structure," Harvard Educational Review 34:246-266.
- Rohrman, N. L. (1968). "The Role of Syntactic Structure in the Recall of English Nominalizations," Journal of Verbal Learning and Verbal Behavior 7:904-912.
- \_\_\_\_\_, D. J. Polzella and L. A. Ackart (1970). "The Psychological Status of Some Syntactic and Semantic Markers," Papers in Linguistics 2:267-278.
- Rosenberg, S. (1966). "Recall of Sentences as a Function of Syntactic and Associative Habit," Journal of Verbal Learning and Verbal Behavior 5:392-396.
- Ross, John R. (1970). "Gapping and the Order of Constituents," in Manfred Bierwisch and Karl Erich Heidolph (eds.), Progress in Linguistics. The Hague: Mouton. Janua Linguarum Series Maior No. 43. Pp. 249-259.
- Saporta, Sol (1967). Psycholinguistic Theories and Generative Grammars. Eugene, Oregon: Oregon State System of Higher Education.



- Savin, H. B. and Ellen Perchonock (1965). "Grammatical Structure and the Immediate Recall of English Sentences," Journal of Verbal Learning and Verbal Behavior 4:348-353.
- Schlesinger, I. M. (1968). Sentence Structure and the Reading Process. The Hague: Mouton. Janua Linguarum Series Minor No. 69.
- Scholes, R. J. (1971). "On the Spoken Disambiguation of Superficially Ambiguous Sentences," Language and Speech 14:1-11.
- Schwartz, D., J. P. Sparkman, and J. Deese (1970). "The Process of Understanding and Judgments of Comprehensibility," Journal of Verbal Learning and Verbal Behavior 9:87-93.
- Smith, Carlota S. (1969). "Ambiguous Sentences with And," in David A. Reibel and Sanford A. Schane (eds.), Modern Studies in English: Readings in Transformational Grammar. Englewood Cliffs, New Jersey: Prentice-Hall. Pp. 75-79.
- Stolz, Walter S. (1967). "A Study of the Ability to Decode Grammatically Novel Sentences," Journal of Verbal Learning and Verbal Behavior 6:867-873.
- \_\_\_\_\_ (1969). "Some Experiments with Queer Sentences," Language and Speech 12:203-219.
- Thorne, J. P. (1966). "On Hearing Sentences," in J. Lyons and R. J. Wales (eds.), Psycholinguistics Papers. Pp. 3-25.
- Tulving, E. and J. E. Patkau (1962). "Concurrent Effects of Contextual Constraint and Word Frequency on Immediate Recall and Learning of Verbal Material," Canadian Journal of Psychology 16:83-95.
- Vanderslice, Ralph (1970). "Occam's Razor and the So-Called Stress Cycle," Language Sciences No. 13, pp. 9-15.
- Wang, M. D. (1970a). "Influence of Linguistic Structure on Comprehensibility and Recognition," Journal of Experimental Psychology 85:83-98.
- \_\_\_\_\_ (1970b). "The Role of Syntactic Complexity as a Determiner of Comprehensibility," Journal of Verbal Learning and Verbal Behavior 9:398-404.

- Wather-Dunn, Weiant (ed.) (1967). Models for the Perception of Speech and Visual Form. Cambridge, Mass.: M.I.T. Press.
- Weinreich, Uriel (1966). "Explorations in Semantic Theory," in Thomas A. Sebeok (ed.), Current Trends in Linguistics, Vol. III, Theoretical Foundations. The Hague: Mouton. Pp. 395-477.
- Wells, Rulon (1947). "Immediate Constituents," Language 23:81-117.
- Wright, Patricia (1969). "Two Studies of the Depth Hypothesis," British Journal of Psychology 60:63-69.
- Yngve, V. A. (1960). "A Model and an Hypothesis for Language Structure," Proceedings of the American Philosophical Society 104:444-466.
- Ziff, Paul (1965). "About What an Adequate Grammar Couldn't Do," Foundations of Language 1:5-13. Reprinted in Jay F. Rosenberg and Charles Travis (eds.), Readings in the Philosophy of Language. Englewood Cliffs, New Jersey: Prentice-Hall. Pp. 548-556. Page references in the text are to the latter source.

APPENDIX A  
Test Sentences\*

\*Sentences so indicated were not used in Experiment II.

Serial Order	Test Form	Surface Structure Sentence Pairs	Percent of Ss Judging Sents. Ambig.	Index of Ambiguity
30	A	The lazy farmer's wife stayed at home.	36.53	-1.44
30	B	The farmer's lazy wife stayed at home.		
54	A	Some more emotional speeches were given there.	61.53	1.69
54	B	Some less emotional speeches were given there.		
41	A	*Police pursued the man with a car.	65.38	2.54
41	B	*Police pursued the man driving a car.		
9	A	The four generals decided on the plane.	63.46	1.90
9	B	The four generals decided in the plane.		
19	A	It's good for them to win this.	55.76	.35
19	B	For them to win this is good.		
7	B	The pretty girl's coat was sent back.	31.48	-2.70
7	A	The coat of the pretty girl was returned.		
18	B	Some more intelligent students were chosen to go.	51.85	.11
18	A	Some less intelligent students were chosen to go.		
31	B	*They discussed the problem with the candidate.	33.33	-2.33
31	A	*They discussed the problem with the campaign.		
43	B	Some ministers laughed out loud at the church.	65.38	2.50
43	A	Some ministers laughed out loud in the church.		
23	B	The foolish boy gave her cat food.	75.47	3.57
23	A	The foolish boy gave him cat food.		

Serial Order	Test Form	Deletion Sentence Pairs	Percent of Ss Judging Sents. Ambig.	Index of Ambiguity
13	A	*Those who play as well as Bill came.	34.61	-1.75
13	B	*Those who play less well than Bill came.		
39	A	John knows a taller man than David.	80.76	4.20
39	B	John knows a taller man than Margaret.		
10	A	He knows more beautiful women than Julia.	75.00	3.39
10	B	He knows women more beautiful than Julia.		
32	A	The visitors saw more than the prisoners.	65.38	1.98
32	B	The visitors saw more than the prisoners saw.		
50	A	Their valuable recordings and films were stolen.	23.07	-3.51
50	B	Their recordings and valuable films were stolen.		
34	B	Lucy likes money more than her husband.	70.37	2.80
34	A	Lucy likes money more than her husband does.		
51	B	Observers thought Sue won and Mary too.	57.40	1.17
51	A	Observers thought Sue won and Mary did too.		
6	B	*Joe watches the fish and also his cat.	38.46	-1.49
6	A	*Joe watches the fish and his cat.		
44	B	They had been arrested and we followed.	64.81	2.02
44	A	They had been arrested and we followed them.		
12	B	He sold his modern paintings and furnishings.	18.51	-4.39
12	A	He sold his paintings and modern furnishings.		

Serial Order	Test Form	Deep Structure Sentence Pairs	Percent of Ss Judging Sents. Ambig.	Index of Ambiguity
42	A	The watching of the spy continued for years.	65.38	2.27
42	B	The wandering of the spy continued for years.		
20	A	The kicking of the horse was futile.	53.84	.66
20	B	The showing of the horse was futile.		
37	A	These friends are the ones to help.	76.92	3.35
37	B	These friends are the ones who help.		
46	A	*The French teacher enjoys sports and music.	26.92	-2.96
46	B	*The chemistry teacher enjoys sports and music.		
33	A	What she feared was being discovered by everyone.	51.92	.68
33	B	What she feared was discovered by everyone else.		
36	B	The testing of the psychologist was revealing.	59.25	1.34
36	A	The traveling of the psychologist was revealing.		
47	B	The painting of the student was realistic.	64.81	2.06
47	A	The painting by the student was realistic.		
25	B	Her special turkey is ready to eat.	44.44	-.70
25	A	Her special turkey is ready to be eaten.		
53	B	*The Spanish scholar wrote a popular book.	14.81	-4.91
53	A	*The young scholar wrote a popular book.		
27	B	What amused them was being ignored by everyone.	53.70	.48
27	A	What amused them was ignored by everyone else.		

Serial Order	Test Form	Case Structure Sentence Pairs	Percent of Ss Judging Sents. Ambig.	Index of Ambiguity
11	A	The very handsome young soldier was crippled.	23.08	-3.50
11	B	The very handsome young soldier was seized.		
29	A	The heavy hammer struck the nail.	17.30	-4.42
29	B	The nail was struck with the heavy hammer.		
48	A	Their big strong brother hit the windshield.	61.53	1.84
48	B	Their big strong ladder hit the windshield.		
55	A	The club president is a beautiful singer.	51.92	.52
55	B	The president of the club sings beautifully.		
35	A	*The green acid quickly dissolved the metal.	9.61	-5.53
35	B	*The metal quickly dissolved in the green acid.		
16	B	Strangers frightened our little black cat.	20.37	-4.15
16	A	Strangers were frightening to our little black cat.		
49	B	The skilled carpenter struck the nail.	20.75	-4.00
49	A	The nail was struck by the skilled carpenter.		
15	B	The happy girl slid across the room.	18.51	-4.32
15	A	The old chair slid across the room.		
22	B	The former peasant was a good king.	16.67	-4.52
22	A	The former peasant was a healthy king.		
21	B	*The fat lady touched the antique clock.	12.96	-5.04
21	A	*The fat lady and the antique clock touched.		

Serial Order	Anomalous Sentences
24	Black heavy dissolved strong the is nail.
26	Our struck the beautifully she chair former.
52	Was chemistry them ready wrote revealing what.
38	Furnishings their more those had and Sue.
40	Them loud pursued home cat generals some.
8	At emotional heavy John for soldier as.
45	Foolish she good them modern young the.
14	Stolen a green the everyone decided teacher.
28	The her skilled the Sue husband were.
17	And prisoners the quickly four beautiful was.



Serial Order

Practice Sentences

---

- |   |  |
|---|--|
| 1 | We found what we were looking for.             |
| 2 | He wanted his brother to wash his car.         |
| 3 | Some politicians refuse to discuss new issues. |
| 4 | Of very the language other to museum.          |
| 5 | Thomas will make a very fine statue.           |

## APPENDIX B

### Instructions to Subjects Judging the Ambiguity of Visually Presented Stimuli

Many English sentences and non-sentences will be projected before you one at a time. Some will be quite simple to understand whereas others will seem quite strange; some may have more than one meaning, and for others it may be impossible to figure out the sense. Each item will be shown only once so please read carefully and try to understand it as you would an ordinary sentence.

After each item there will be a brief pause during which you are asked to do two things. First, indicate on the answer sheet by underlining the appropriate word whether you think the item you just read is ambiguous, i.e. has more than one possible meaning, is unambiguous, i.e. has a single meaning, or is meaningless. This judgment should be based upon your overall reaction to the entire segment, not just a part of it. Following each set of words on the answer sheet you will see the numbers 1 through 7. If you are very sure of the judgment you just indicated, circle the 7. If you are less confident of your judgment but think it is probably accurate, circle the 6. If you are still less sure, circle the 5. If you think your chance of being in error is about 50-50, then circle the 4. If you are distinctly uncertain about this judgment, circle the 3; very uncertain, the 2. Circle 1 if you have no confidence in your judgment. To repeat, circling the 7 means you are very sure of yourself, the 1 very unsure.

We are interested in your initial reaction to each item so please watch closely and respond quickly and frankly. Your identity will not be revealed nor will you be graded in any way on your performance.

Are there any questions?

Instructions to Subjects Judging the Ambiguity  
of Aurally Presented Stimuli

You are going to hear a tape recording of many English sentences and non-sentences. Some will be quite simple to understand whereas others will seem quite strange; some may have more than one meaning, and for others it may be impossible to figure out the sense. Each item will be played only once so please listen carefully and try to understand it as you would an ordinary sentence.

After each item there will be a brief pause during which you are asked to do two things. First, indicate on the answer sheet by underlining the appropriate word whether you think the item you just heard is ambiguous, i.e. has more than one possible meaning, is unambiguous, i.e. has a single meaning, or is meaningless. This judgment should be based upon your overall reaction to the entire segment, not just a part of it. Following each set of words on the answer sheet you will see the numbers 1 through 7. If you are very sure of the judgment you just indicated, circle the 7. If you are less confident of your judgment but think it is probably accurate, circle the 6. If you are still less sure, circle the 5. If you think your chance of being in error is about 50-50, then circle the 4. If you are distinctly uncertain about this judgment, circle the 3; very uncertain, the 2. Circle 1 if you have no confidence in your judgment. To repeat, circling the 7 means you are very sure of yourself, the 1 very unsure.

We are interested in your initial reaction to each item so please listen closely and respond quickly and frankly. Your identity will not be revealed nor will you be graded in any way on your performance.

Are there any questions?

Instructions to Subjects Judging the Comprehensibility  
of Visually Presented Stimuli

Many English sentences and non-sentences will be projected before you one at a time. Some will be quite simple to understand whereas others will seem quite strange; some may have more than one meaning, and for others it may be impossible to figure out the sense. Each item will be shown only once so please read carefully and try to understand it as you would an ordinary sentence.

After each item there will be a brief pause during which you are asked to do two things. First, indicate on the answer sheet by underlining the appropriate word whether you think the item you just read is comprehensible, i.e. easy to understand, or incomprehensible, i.e. impossible to understand. This judgment should be based upon your overall reaction to the entire segment, not just a part of it. Following each pair of words on the answer sheet you will see the numbers 1 through 7. If you are very sure of the judgment you just indicated, circle the 7. If you are less confident of your judgment but think it is probably accurate, circle the 6. If you are still less sure, circle the 5. If you think your chance of being in error is about 50-50, then circle the 4. If you are distinctly uncertain about this judgment, circle the 3; very uncertain, the 2. Circle 1 if you have no confidence in your judgment. To repeat, circling the 7 means you are very sure of yourself, the 1 very unsure.

We are interested in your initial reaction to each item so please watch closely and respond quickly and frankly. Your identity will not be revealed nor will you be graded in any way on your performance.

Are there any questions?

Instructions to Subjects Judging the Comprehensibility  
of Aurally Presented Stimuli

You are going to hear a tape recording of many English sentences and non-sentences. Some will be quite simple to understand whereas others will seem quite strange; some may have more than one meaning, and for others it may be impossible to figure out the sense. Each item will be played only once so please listen carefully and try to understand it as you would an ordinary sentence.

After each item there will be a brief pause during which you are asked to do two things. First, indicate on the answer sheet by underlining the appropriate word whether you think the item you just heard is comprehensible, i.e. easy to understand, or incomprehensible, i.e. impossible to understand. This judgment should be based upon your overall reaction to the entire segment, not just a part of it. Following each pair of words on the answer sheet you will see the numbers 1 through 7. If you are very sure of the judgment you just indicated, circle the 7. If you are less confident of your judgment but think it is probably accurate, circle the 6. If you are still less sure, circle the 5. If you think your chance of being in error is about 50-50, then circle the 4. If you are distinctly uncertain about this judgment, circle the 3; very uncertain, the 2. Circle 1 if you have no confidence in your judgment. To repeat, circling the 7 means you are very sure of yourself, the 1 very unsure.

We are interested in your initial reaction to each item so please listen closely and respond quickly and frankly. Your identity will not be revealed nor will you be graded in any way on your performance.

Are there any questions?

Instructions to Subjects Judging the Complexity  
of Visually Presented Stimuli

Many English sentences and non-sentences will be projected before you one at a time. Some will be quite simple to understand whereas others will seem quite strange; some may have more than one meaning, and for others it may be impossible to figure out the sense. Each item will be shown only once so please read carefully and try to understand it as you would an ordinary sentence.

After each item there will be a brief pause during which you are asked to do two things. First, indicate on the answer sheet by underlining the appropriate word(s) whether you think the item you just read is complex, i.e. complicated, not in the grammatical sense of the term, or is simple, i.e. uncomplicated or straightforward. This judgment should be based upon your overall reaction to the entire segment, not just a part of it. Following each pair of terms on the answer sheet you will see the numbers 1 through 7. If you are very sure of the judgment you just indicated, circle the 7. If you are less confident of your judgment but think it is probably accurate, circle the 6. If you are still less sure, circle the 5. If you think your chance of being in error is about 50-50, then circle the 4. If you are distinctly uncertain about this judgment, circle the 3; very uncertain, the 2. Circle 1 if you have no confidence in your judgment. To repeat, circling the 7 means you are very sure of yourself, the 1 very unsure.

We are interested in your initial reaction to each item so please watch closely and respond quickly and frankly. Your identity will not be revealed nor will you be graded in any way on your performance.

Are there any questions?

Instructions to Subjects Judging the Complexity  
of Aurally Presented Stimuli

You are going to hear a tape recording of many English sentences and non-sentences. Some will be quite simple to understand whereas others will seem quite strange; some may have more than one meaning, and for others it may be impossible to figure out the sense. Each item will be played only once so please listen carefully and try to understand it as you would an ordinary sentence.

After each item there will be a brief pause during which you are asked to do two things. First, indicate on the answer sheet by underlining the appropriate word(s) whether you think the item you just heard is complex, i.e. complicated, not in the grammatical sense of the term, or is simple, i.e. uncomplicated or straightforward. This judgment should be based upon your overall reaction to the entire segment, not just a part of it. Following each pair of terms on the answer sheet you will see the numbers 1 through 7. If you are very sure of the judgment you just indicated, circle the 7. If you are less confident of your judgment but think it is probably accurate, circle the 6. If you are still less sure, circle the 5. If you think your chance of being in error is about 50-50, then circle the 4. If you are distinctly uncertain about this judgment, circle the 3; very uncertain, the 2. Circle 1 if you have no confidence in your judgment. To repeat, circling the 7 means you are very sure of yourself, the 1 very unsure.

We are interested in your initial reaction to each item so please listen closely and respond quickly and frankly. Your identity will not be revealed nor will you be graded in any way on your performance.

Are there any questions?

## APPENDIX C

## Answer Sheet Format for Ambiguity Judgments

1. Ambiguous: Unambiguous: Meaningless

Extremely Confident	(7)	6	5	4	3	2	1	No Con- fidence
------------------------	-----	---	---	---	---	---	---	--------------------

2. Ambiguous: Unambiguous: Meaningless

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

3. Ambiguous: Unambiguous: Meaningless

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

4. Ambiguous: Unambiguous: Meaningless

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

5. Ambiguous: Unambiguous: Meaningless

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

6. Ambiguous: Unambiguous: Meaningless

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

7. Ambiguous: Unambiguous: Meaningless

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

8. Ambiguous: Unambiguous: Meaningless

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------



# Answer Sheet Format for Comprehensibility Judgments

## 1. Comprehensible:Incomprehensible

Extremely Confident	(7)	6	5	4	3	2	1	No Con- fidence
------------------------	-----	---	---	---	---	---	---	--------------------

## 2. Comprehensible:Incomprehensible

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

## 3. Comprehensible:Incomprehensible

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

## 4. Comprehensible:Incomprehensible

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

## 5. Comprehensible:Incomprehensible

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

## 6. Comprehensible:Incomprehensible

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

## 7. Comprehensible:Incomprehensible

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

## 8. Comprehensible:Incomprehensible

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

### Answer Sheet Format for Complexity Judgments

1. Complex:Not Complex

Extremely Confident	(7)	6	5	4	3	2	1	No Con- fidence
------------------------	-----	---	---	---	---	---	---	--------------------

2. Complex:Not Complex

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

3. Complex:Not Complex

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

4. Complex:Not Complex

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

5. Complex:Not Complex

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

6. Complex:Not Complex

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

7. Complex:Not Complex

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

8. Complex:Not Complex

Extremely Confident	7	6	5	4	3	2	1	No Con- fidence
------------------------	---	---	---	---	---	---	---	--------------------

## APPENDIX D

### Instructions Read to Subjects in Experiment II

Please arrange the pad of answer sheets so that the glued edges are on the left side and you can leaf through the pages as if you were reading a book. On the first page write your name, your class standing, and your major.

On the tape I'm about to play you will hear a number of English sentences which have been recorded by one of my male colleagues. After each sentence there is a brief pause before the narrator pronounces seven pairs of digits. These numbers are ordered so that the first pair is always somewhere in the range of 20 to 29, the second pair ranges from 30 to 39, the third pair from 40 to 49, and so on through the final pair which ranges from 80 to 89. Your task is to listen carefully and after you hear the instruction "begin" you are to write the sentence across the top of the page exactly as you remember it and then write down as many numbers as possible from the list. The numbers may be written down in any order, that is, not necessarily in the same order in which you heard them. The sentence, however, must be completed before you begin the numbers. Remember, do not start writing until you hear the word "begin."

When you finish writing, turn the page and get ready for the next item. There are 37 items on this test and I will stop the recorder for brief rest periods after 12 or 13 sentences.

Are there any questions?

### After the Test

Now go back through your answers and underline those sentences which you think are ambiguous, i.e. have more than one meaning.

## VITA

Carol June Cunningham was born 3 June 1941 in Chicago, Illinois. She attended public school in Westmont, Illinois, and graduated as valedictorian of Hinsdale Township High School in 1958. Four years later she received a Bachelor of Arts degree from the University of Colorado where she was elected to membership in Phi Beta Kappa. From 1962 until 1967 she was employed as a technical translator by the Tennessee Valley Authority, National Fertilizer Development Center, Muscle Shoals, Alabama. In 1964 she married Dr. James F. Parr, Jr. Their daughter, Lauren Melissa, was born in Baton Rouge, 29 August 1968.

Ms. Parr was awarded a Master of Arts degree by Louisiana State University in 1970. During the period of her graduate education she has expressed her concern about injustices in women's roles in American society by participating in numerous activities of the National Organization for Women.

## EXAMINATION AND THESIS REPORT

Candidate: Carol Cunningham Parr

Major Field: Linguistics

Title of Thesis: The Effect of Syntactic Ambiguity on Judgments of the Comprehensibility and Complexity of Sentences and on Short-Term Memory Capacity

Approved:

*J. Donald Ragsdale*  
Major Professor and Chairman

*Max Goodrich*  
Dean of the Graduate School

### EXAMINING COMMITTEE:

*Mary L. Metz*

*C. L. Shaver*

*Margaret S. Stanley*

*William W. Evans*

Date of Examination:

5 July 1972